1. Testing hypotheses about means. (PCN 6.6)
Testing hypotheses about the mean of a single population (PCN 6.6)

Examples of where this might be interesting:

- Comparison to a standard (Example 6.4: Post-partum hospital stays)
- Testing a manufacturer’s claim (e.g., Net weight: 1.75 oz.)

General idea:

1. Take a sample from the population you want to make inference about.
2. Compute the sample mean.
3. Compare the sample mean to the hypothesized mean (i.e., the standard or claim)
4. Make a statement about how likely you would have been to get the value for the sample mean that you observed if the hypothesized mean is true.
5. Draw a conclusion about the mean of your population relative to the hypothesized mean.
Testing the statistical hypothesis:

The scientific hypothesis

The statistical model

- The statistical model describes the population you are sampling from.
- Here we assume that our measurements follow a normal distribution with (unknown) mean $\mu$ and variance $\sigma^2$.

$$Y_1, Y_2, \ldots, Y_n \sim N(\mu, \sigma^2)$$

The statistical hypothesis

- Here we make some statement about the mean of our target population (i.e., $\mu$) relative to a standard or claimed mean that we will call $\mu_0$.
- In most situations $\mu_0$ is a number (e.g., $\mu_0 = 1.75$), but we leave it as a symbol here for generality.
- The null hypothesis, $H_0$, represents the case where the mean of the target population is equal to the standard or claimed mean.

$$H_0 : \mu = \mu_0$$

- The alternative hypothesis, $H_a$, represents the case where our scientific hypothesis is true. For this there are three possibilities:

$$H_a : \mu < \mu_0 \quad H_a : \mu \neq \mu_0 \quad H_a : \mu > \mu_0$$
The test statistic

- Your book defines a test statistic as a function of the data that does not involve any of the parameters. This is the correct definition. However...

- If $\sigma^2$ is known, we will define the test statistic for any of
  \[
  H_0 : \mu = \mu_0 \quad H_0 : \mu = \mu_0 \quad H_0 : \mu = \mu_0 \\
  H_a : \mu < \mu_0 \quad H_a : \mu \neq \mu_0 \quad H_a : \mu > \mu_0
  \]
  to be
  \[
  z^* = \frac{\bar{Y} - \mu_0}{\sigma/\sqrt{n}}
  \]

- If $\sigma^2$ is not known then the test statistic is defined to be
  \[
  t^* = \frac{\bar{Y} - \mu_0}{s/\sqrt{n}}
  \]
  where
  \[
  s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2}
  \]
  is the estimate of the population standard deviation $\sigma$ based on our sample.

- Why the two cases?
  1. If we don’t know the value of $\sigma^2$ then we can’t compute $z^*$.
  2. If we do know the value of $\sigma^2$ then we should use that rather than an estimate (i.e., $S$).
Example of what we’ve talked about so far today.

1. Professor Swift routinely buys Doritos from vending machines on campus. He notices that the manufacturer claims that the net weight of chips in each bag is 1.75 ounces, but he thinks it must be much lower. What is the “scientific” hypothesis? (Note: This is a quality control problem.)

2. To test this hypothesis, Professor Swift starts weighing the chips in each bag before eating them. What is the statistical model?

3. What is the statistical hypothesis?

4. Over the course of two weeks he “samples” 9 bags and computes the following statistics.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$Y$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1.65</td>
<td>0.1</td>
</tr>
</tbody>
</table>

   What is the appropriate test statistic?
Comparing the observed test statistic to the sampling distribution

Recall:

- We test the hypothesis under the assumption that $H_0$ is true.
- We use the sampling distribution of our test statistic to determine the plausibility of $H_0$, in light of the data we observed.

What are the sampling distributions of our test statistics?

- If $H_0$ is true, $z^*$ follows a $N(0, 1)$ distribution.
- If $H_0$ is true, $t^*$ follows a $t_{n-1}$ distribution.
- How did we get these distributions?
  (See MA 3631 or MA 4632 for theory.)
- Why do they have different sampling distributions?
How can we use this information to make a decision about our hypothesis?

There are two equivalent methods

- the \( p \)-value approach.
- the Critical Value approach.

The \( p \)-value approach to testing hypotheses

Compute the \( p \)-value and compare it to the pre-specified level of significance \( \alpha \). (e.g., \( \alpha = 0.05 \)). But how do we compute the \( p \)-value?

The \( p \)-value is always the probability of observing a test statistic as extreme (or more extreme) than you did if the null hypothesis were true.

- For \( z^* \) you can use Table A.3 in the back of your book.
  1. For \( H_a : \mu < \mu_0 \) the \( p \)-value is the probability that a \( N(0,1) \) random variable is less than or equal to \( z^* \): \( P(Z \leq z^*) \)
  2. For \( H_a : \mu > \mu_0 \) the \( p \)-value is the probability that a \( N(0,1) \) random variable is greater than or equal to \( z^* \):
     \[ P(Z \geq z^*) = 1 - P(Z \leq z^*) \]
  3. For \( H_a : \mu \neq \mu_0 \) the \( p \)-value is the probability that a \( N(0,1) \) random variable is greater than or equal to \( z^* \) in absolute value:
     \[ P(|Z| \geq z^*) = P(Z \leq -|z^*|) + P(Z \geq |z^*|) = 2P(Z \leq -|z^*|) \]

- I have defined all these \( p \)-values in terms of \( P(Z \leq z) \) because those are the values that are given in Table A.3.
  Warning: Different textbooks use different table formats.

- For \( t^* \) you can only get a range for the \( p \)-value using Table A.4 in the back of your book.

- SAS will also compute \( p \)-values.
What if our test statistic is $t^*$?

- If $H_0$ is true, $t^*$ follows a $t_{n-1}$ distribution. (Table A.4)
- We can’t find an exact $p$-value using Table A.4.
- We can find a range for the $p$-value. Sometimes this is enough.

How to find a range for the $p$-value associated with $t^*$ when testing a hypothesis about a population mean:

1. Compute the degrees of freedom (df). For tests about the mean of a normal population $df = n - 1$, one less than the number of samples.
2. Go to the row in Table A.4 with the appropriate degrees of freedom.
3. For $H_a : \mu < \mu_0$ the $p$-value is the probability that a $t_{n-1}$ random variable is less than or equal to $t^*$: $P(t_{n-1} \leq t^*)$
   - If $t^* < 0$, which it usually will be for this hypothesis...
     - Read across the row until you find the two numbers between which $-t^*$ falls. For example, if $df = 8$ and $t^* = -3$, $-t^* = 3$ falls in between 2.8965 and 3.3554.
     - Go up to the top of the column and see which quantiles of the $t_{n-1}$ distribution these two numbers are. Let’s call them $q_1$ and $q_2$. For example, 2.8965 is the 0.99 quantile of the $t_8$ distribution and 3.3554 is the 0.995 quantile of the $t_8$ distribution. In this case $q_1 = 0.99$ and $q_2 = 0.995$.
     - The $p$-value is in the range $(1 - q_2, 1 - q_1)$. For our example, the $p$-value is in the range $(1 - 0.995, 1 - 0.99) = (0.005, 0.01)$
     - Exception 1: If $-t^* < t_{df,0.90}$, the $p$-value is in the range $(0.10, 0.50)$
     - Exception 2: If $-t^* > t_{df,0.9995}$, the $p$-value is in the range $(0, 0.0005)$
• If \( t^* = 0 \), which is extremely rare in practice, the \( p \)-value is 0.50.

• If \( t^* > 0 \), which rarely will be for this hypothesis...
  
  - Read across the row until you find the two numbers between which \( t^* \) falls. For example, if \( df = 10 \) and \( t^* = 2 \), \( t^* = 2 \) falls in between 1.8125 and 2.2281.

  - Go up to the top of the column and see which quantiles of the \( t_{n-1} \) distribution these two numbers are. Let’s again call them \( q_1 \) and \( q_2 \). For example, 1.8125 is the 0.95 quantile of the \( t_{10} \) distribution and 2.2281 is the 0.975 quantile of the \( t_{10} \) distribution. In this case \( q_1 = 0.95 \) and \( q_2 = 0.975 \).

  - The \( p \)-value is in the range \((q_1, q_2)\). For our example, the \( p \)-value is in the range \((0.95, 0.975)\)

  - Exception 1: If \( t^* \leq t_{df, 0.90} \), the \( p \)-value is in the range \((0.50, 0.90)\)

  - Exception 2: If \( t^* \geq t_{df, 0.9995} \), the \( p \)-value is in the range \((0.9995, 1)\)

4. Draw a conclusion.

  • If the entire range for the \( p \)-value is less than \( \alpha \), reject \( H_0 \).

  • If the entire range for the \( p \)-value is greater than \( \alpha \), do not reject \( H_0 \).

  • If \( \alpha \) falls in the range for the \( p \)-value you cannot draw any conclusion.

5. For \( H_a : \mu > \mu_0 \) the \( p \)-value is the probability that a \( t_{n-1} \) random variable is greater than or equal to \( t^* \):

\[
P(t_{n-1} \geq t^*) = 1 - P(t_{n-1} \leq t^*) = P(t_{n-1} \leq -t^*)
\]

6. For \( H_a : \mu \neq \mu_0 \) the \( p \)-value is the probability that a \( t_{n-1} \) random variable is greater than or equal to \( t^* \) in absolute value:

\[
P(|t_{n-1}| \geq t^*) = P(t_{n-1} \leq -|t^*|) + P(t_{n-1} \geq |t^*|) = 2P(t_{n-1} \leq -|t^*|)
\]

An equally complicated set of rules could be constructed for computing \( p \)-values for these hypotheses, but since this is ridiculously complicated, there is another method for drawing conclusions about statistical hypotheses.
In this second **equivalent** method we reach our conclusion about $H_0$ based on what is referred to as a **critical value**. (PNC 6.10)

If the observed test statistic, $t^*$, provides more evidence against $H_0$ (and in favor of $H_a$) than the critical value for your hypothesis, $H_0$ is rejected. Otherwise, we do not reject $H_0$.

**What is the critical value?**

- The critical value, $t_{crit}$, is the value of $t^*$ that would have produced a $p$-value of $\alpha$.
- Since we know the value of $t_{crit}$ we can tell whether the $p$-value is less than or greater than $\alpha$ and draw a conclusion about $H_0$.

**How do we get the critical value?**

1. Specify the significance level of your test. Usually $\alpha = 0.05$.

2. Find the degrees of freedom for your test statistic. When using $t^*$ to test hypothesis about the mean of a single population, $df = n - 1$.

3. Use Table A.4 to find $t_{crit}$ and compare it to $t^*$.

   (a) For $H_a : \mu < \mu_0$, $t_{crit} = -t_{df, 1-\alpha}$ and if $t^* < t_{crit}$ reject $H_0 : \mu = \mu_0$ in favor of $H_a$. Otherwise do not reject $H_0$.

   (b) For $H_a : \mu > \mu_0$, $t_{crit} = t_{df, 1-\alpha}$ and if $t^* > t_{crit}$ reject $H_0 : \mu = \mu_0$ in favor of $H_a$. Otherwise do not reject $H_0$.

   (c) For $H_a : \mu \neq \mu_0$, $t_{crit} = t_{df, 1-\alpha/2}$ and if $|t^*| > t_{crit}$ reject $H_0 : \mu = \mu_0$ in favor of $H_a$. Otherwise do not reject $H_0$.

**Example:** Doritos

<table>
<thead>
<tr>
<th>$H_0$ : $\mu = 1.75$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_a$ : $\mu &lt; 1.75$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$n$</th>
<th>$Y$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>1.65</td>
<td>0.1</td>
</tr>
</tbody>
</table>