Example 1  (Lecture 16)

We want to test the hypotheses:

\[ H_0 : \mu = 50 \]
\[ H_A : \mu < 50 \]

Let \( X_i \) be the weight of a randomly chosen chocolate with mean \( \mu \) and variance \( \sigma^2 \). We evaluate the sample mean \( \bar{X} \) based on a sample of size 16.

We suppose that the CLT has kicked in and

\[ \bar{X} \sim N \left( \mu, \frac{\sigma^2}{n} \right) \]

This is the quantity we want to test.

We need to look at \( \bar{X} \) under the null. Because the mean \( \mu = 50 \), so too "big" we only look at the distribution when \( \mu = 50 \).

When \( \mu = 50 \) the distribution of the sample mean looks like
Distributon $\bar{X}$ under the assumption

$\bar{X} \sim N(50, \frac{2}{16})$.

Rejection region should not contain the mean under the null hypothesis $\mu = 50$.

$50 - 1.64 \sqrt{\frac{2}{16}} = 49.4$

When $\bar{X}$ lies in this region here we reject the null at the 5% level.

$\cdot$ This is the rejection region method for doing hypothesis tests. Another equivalent method is to look at the p-value. P-values are usually given in computer output, so it is a good idea to understand this method, but the rejection region method really helps understanding too.
To calculate the p-value you need to calculate the probability of a certain region under the null.

Typically, this region should start at the observed sample mean \( \bar{x} = 47.8 \) and go to either \(-\infty\) or \(+\infty\). You choose the interval (region) which does not contain the \( \mu \) under the null.

Since in this example \( H_0: \mu \geq 50 \)

The interval \( [47.8, 50] \) contains all points greater than 50 or equal to 50. We cannot use this interval. We choose instead the interval...
This interval does not contain 50 and above.

Therefore we calculate the probability \( \bar{X} \) lies between \(-\infty\) and 47.8 given that \( \bar{X} \sim N(50, \frac{2}{16}) \).

The p-value is

\[
P\left( \bar{X} \leq 47.8 \mid \bar{X} \sim N(50, \frac{2}{16}) \right)
\]

This probability can be calculated using the usual standardisation.
Look up & Tables. You will see that

$$P\left\{ \bar{X} \leq 47.8 \mid \bar{X} \sim N\left(50, \frac{2}{16}\right) \right\} = \text{very small [0.0001]}.$$  

Since this probability is so small and less than 5%, there is evidence to reject the null and suppose the mean weight of the chocolate bar is less than 50 grams.
Remember that the conjecture you want to check so that the mean is not 50 grams. So state this as the alternative:

\[ H_0: \mu = 50 \quad [\text{mean is 50 grams}] \]

\[ H_A: \mu \neq 50. \]

Previously (Example 1) was a one-sided test. This time we are doing a 2-sided test. But everything is the same as before just the rejection region on significance level has changed.

Under the null \( \overline{X} \sim N(50, \frac{2}{16}) \).

Because we are interested in checking if \( \mu \neq 50 \), if \( \overline{X} \) is either too big or too small, there is enough evidence to reject the null. Too big or small, means that that the rejection region is on both sides of the mean.
(a) Because it's a 2-sided test the sum of the areas on both sides should add to 5%. That is why it is 2.5% on either side.

(b) To evaluate the p-value we choose the region under which does not contain the mean under the null and evaluate the probability in that region.

\[ P \{ \bar{X} \leq 47.8 \mid \bar{X} \sim N(50, \frac{2}{16}) \} \]

This is exactly the same probability as before, \( = 0.0001 \) (very small).
Because it is a 2-sided test we compare

$$P \left[ \bar{X} \leq 47.8 \mid \bar{X} \sim N(50, \frac{2}{16}) \right]$$

with 2.5% (not 5%). If this probability

is less than 2.5%, then there is evidence
to reject the null.

Since

$$P \left[ \bar{X} \leq 47.8 \mid \bar{X} \sim N(50, \frac{2}{16}) \right] \leq 0.025$$

there is evidence to reject the null and suppose
the mean is not 50.
Solution 3

We remember the lawyer wants to test whether the mean is greater than 50.

\[ H_0: \mu \leq 50 \]
\[ H_A: \mu > 50 \]

This is a one-sided test.

Under the null \( \bar{X} \sim N \left( 50, \frac{2}{16} \right) \) or less.

Since \( \bar{X} \) under the null has a mean of 50 or less, we only look at the case \( \bar{X} \sim N \left( 50, \frac{2}{16} \right) \).

The rejection region should not contain the mean under the null (which is \( \mu \leq 50 \)).

\[
5_{\%} \quad 50 \quad 50 + 1.645 \sqrt{\frac{2}{16}} = 50.6
\]
We observe that $\bar{X} = 47.8$. Clearly this is not in the rejection region. Therefore there is not enough evidence to reject the null.

It is clear from the sample mean $\bar{X} = 47.8$, that we would not be able to reject the null. This is because the sample mean lies in the null hypotheses region.

It is like saying that you are sure the mean should be at least greater than 50, but all the evidence you have is a sample mean of 47.8! This statement has no meaning.

In this case the p-value has no meaning. Both the intervals

\[ -\infty \quad 47.8 \quad 47.8 \quad \text{infty} \]

contain values in the null region ($\mu \leq 50$). Here the p-value cannot be evaluated.
In the case of a one-sided test if \( \bar{X} \) lies in \( H_0 \) (so you see that 47.8 \( \leq \) 50 - it lies in \( H_0 \)), then we cannot reject the null.
Types of Tests

a) Two-sided Tests  Thus we when in the null the parameters is equal to some value.

\[ H_0: \mu = \mu_0 \quad (\text{say } \mu = 5) \]
\[ H_A: \mu \neq \mu_0 \quad (\text{say } \mu \neq 5) \]

The distribution under the null looks like:

(a) If p-value is less than 2.5%, we reject the null.

\[ \bar{x} \]
\[ \mu_0 (5) \]

\[ \uparrow \text{The average we observe} \]

\[ \uparrow \text{This area is the p-value (the region should not include } \mu_0 \text{).} \]

(b) If \( \bar{x} \) lies in the rejection region we reject the null.
b) One-sided Test

\[ H_0 : \mu \geq \mu_0 \quad (\mu \geq 5) \]

\[ H_A : \mu < \mu_0 \quad (\mu < 5) \]

Suppose we do the test at the 5% level.

We choose the null hypothesis to look at the distribution when \[ \mu = \mu_0 \]

\[
\begin{array}{c}
5\% \\
\hline
\end{array}
\]

rejection region.

If \( \bar{x} > \mu_0 \), there is not enough evidence to reject the null.

If \( \bar{x} < \mu_0 \), calculate the p-value or check to see whether \( \bar{x} \) lies in the rejection region.

If p-value less than 5% reject null
c) One-sided test

\[ H_0: \mu \leq \mu_0 \quad (\mu \leq 5) \]
\[ H_A: \mu > \mu_0 \quad (\mu > 5) \]

Suppose we do the test at the 5% level.

We look at the distribution when \( \mu = \mu_0 \).

If \( \bar{X} < \mu_0 \), there is not enough evidence to reject the null.

If \( \bar{X} > \mu_0 \), calculate the p-value or check to see whether \( \bar{X} \) lies in the rejection region.

If p-value is greater than or less than 5%, reject null.