Lecture 8:
Hypothesis Testing
STAT 630, Fall 2021

Hypothesis Test for One Mean

Key components:

- Null hypothesis:
  $H_0 : \mu = \mu_0$

- Alternative hypothesis (use one of these):
  $H_A : \mu > \mu_0$ (one-sided, upper-tail)
  $H_A : \mu < \mu_0$ (one-sided, lower-tail)
  $H_A : \mu \neq \mu_0$ (two-sided)

- Test statistic:

$$t = \frac{\text{observed value} - \text{null value}}{\text{SE}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- A rule to either reject or not reject H_0 (based on p-value)

Hypothesis Testing Concept

The approach to hypothesis testing is as follows:

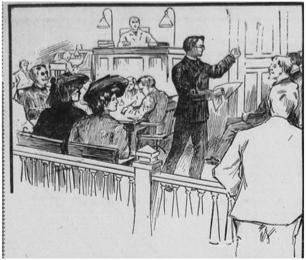
1. Assume that H_0 is true. H_0 usually represents a skeptical position, or a perspective of no difference or change in the parameter of interest.
2. Reject H_0 only if the data provide strong evidence in support of the alternative claim in H_A.

Hypothesis Testing Concept

The hypothesis testing framework can be found in the US court system, where innocence is assumed until proven guilty.[1]

H_0: The defendant is innocent.
H_A: The defendant is guilty.

The jurors consider whether the evidence is convincing enough to convict the defendant (reject H_0).



[1] https://commons.wikimedia.org/wiki/File:Trial_of_Edward_Ellis_(courtroom_sketch).jpg

p-value

A p-**value** is the probability of obtaining a test statistic as extreme, or more extreme (in the direction of the alternative), than the observed value of the test statistic, assuming that H_0 is true.

p-value

Decision rule using the p-value:

- If p-value $< \alpha$, then reject H_0.
- If p-value $> \alpha$, then do not reject H_0.

α is called the **signficance level**.
Common values for $\alpha = 0.05, 0.01$

p-value

- When the p-value $< \alpha$ (we reject H_0) the result is said to be **statistically significant**.

- In other words, a result is statistically significant when it is unlikely to of occurred by random chance, assuming that the null hypothesis is true.

- The smaller the p-value, the stronger the data favor H_A over H_0.

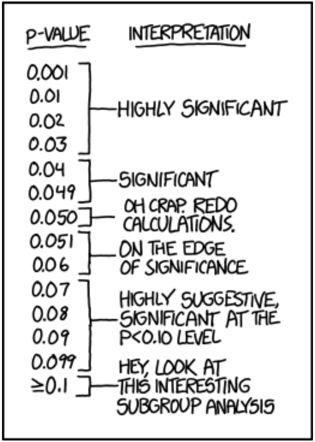


Image from https://xkcd.com/1478/

Computing p-values

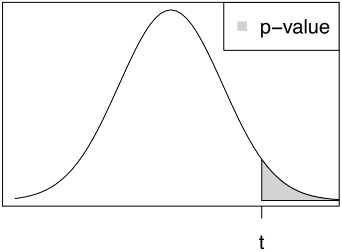
One-sided test (upper-tail):
$H_0 : \mu = \mu_0$
$H_A : \mu > \mu_0$

Test statistic:

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

p-value $=$ `1 - pt(t, df = n-1)`

Reject H_0 if p-value $< \alpha$



Computing p-values

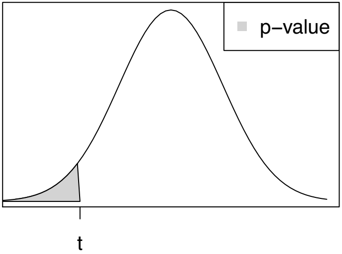
One-sided test (lower-tail):
$H_0 : \mu = \mu_0$
$H_A : \mu < \mu_0$

Test statistic:

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

p-value $=$ `pt(t, df = n-1)`

Reject H_0 if p-value $< \alpha$



Computing p-values

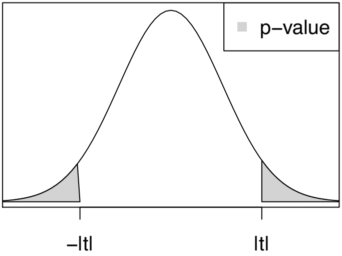
Two-sided test:
$H_0 : \mu = \mu_0$
$H_A : \mu \neq \mu_0$

Test statistic:

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

p-value $=$ `2*pt(-abs(t), df = n-1)`

Reject H_0 if p-value $< \alpha$



Conditions

The hypothesis test is valid when the following conditions are satisfied:

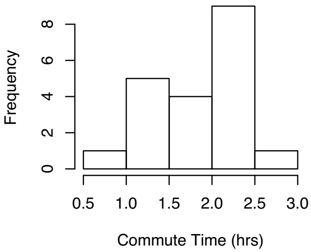
- Sample observations are independent. Generally, this is satisfied when the data come from a random sample.
- The sample size is large ($n \geq 30$), and there are no extreme outliers. This implies that the sampling distribution for $\bar{X}$ is approximately normal according to the central limit theorem.
- Otherwise, if the sample size is small ($n < 30$), the data should follow an approximate normal distribution. Graphical methods can be used to check this (box plot, histogram, normal QQ plot).

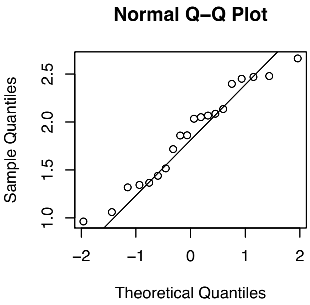
These are the same conditions that we need to check for a confidence interval for the population mean.

Note that when the sample size is large ($n \geq 30$) we can use either a t or z test statistic when computing the p-value (i.e., `pt()` and `pnorm()` will give approximately the same p-value).

Example 1

An administrator is interested in testing whether the average amount of time CSUEB students spend commuting to campus exceeds 1.5 hours. A random sample of $n = 20$ students are interviewed. The sample mean $\bar{x} = 1.86$ and standard deviation $s = 0.5$. A histogram and QQ plot of the data are shown below.





Example 1

(a) Write the null and alternative hypothesis for a one-sided test.

$H_0 : \mu = 1.5$
$H_A : \mu > 1.5$

(b) Check the conditions for the hypothesis test.

The conditions for the test are satisfied:

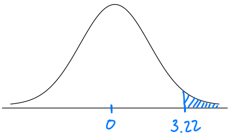
- Data come from a random sample
- Data follow approximate normal distribution (need to check since $n = 20 < 30$)

(c) Calculate the test statistic.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{1.86 - 1.5}{0.5/\sqrt{20}} = 3.22$$

Example 1

(d) Calculate the p-value and make a decision using $\alpha = 0.05$ significance level.



p-value $=$ `1 - pt(3.22, df=19)` $= 0.0023$
Since p-value < 0.05, we reject H_0

(e) What is the conclusion of the test in the context of the data?

The data provide strong evidence that the mean commute time for all CSUEB students is greater than 1.5 hours.

Example 2

It is claimed that the mean mileage of a certain type of vehicle is 35 miles per gallon (mpg) of gasoline. A random sample of 49 vehicles are collected. The sample mean $\bar{x} = 36.1$ mpg and standard deviation $s = 5.4$ mpg.

(a) Write the null and alternative hypothesis for a two-sided test.

$H_0 : \mu = 35$
$H_A : \mu \neq 35$

(b) Check the conditions for the hypothesis test.

The conditions for the test are satisfied:

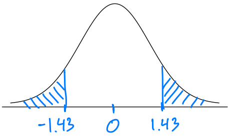
- Data come from a random sample
- Large sample size $n = 49 > 30$

Example 2

(c) Calculate the test statistic.

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{36.1 - 35}{5.4/\sqrt{49}} = 1.43$$

(d) Calculate the p-value and make a decision using $\alpha = 0.05$ significance level.



p-value $=$ `2*pt(-1.43, df=48)` $= 0.159$
Since p-value > 0.05, we do not reject H_0

(e) What is the conclusion of the test in the context of the data?

The population mean mileage is **not** significantly different than 35 mpg.

Decision Errors

Decision \ Truth	H_0 true	H_A true
Reject H_0	Type 1 error (α)	✓
Do not reject H_0	✓	Type 2 error (β)

The significance level α is the probability of a type 1 error:

$$\begin{aligned}\alpha &= \text{Probability of rejecting } H_0\text{, when } H_0 \text{ is true} \\ &= P(\text{reject } H_0 | H_0 \text{ true})\end{aligned}$$

Decision Errors

In a US court, the defendant is either innocent (H_0) or guilty (H_A). What does a type 1 error represent in this context? What does a type 2 error represent?

H_0 : The defendant is innocent
H_A : The defendant is guilty

Type 1 error:
The defendant is actually innocent, but the jury decides guilty.

Type 2 error:
The defendant is actually guilty, but the jury decides innocent.

Using a confidence interval to perform a two-sided hypothesis test

Suppose we want to perform the following two-sided test at the α level of significance:

$H_0 : \mu = \mu_0$
$H_A : \mu \neq \mu_0$

Construct a $1 - \alpha$ confidence interval (CI):

$$\bar{x} \pm t_{\alpha/2;n-1} \frac{s}{\sqrt{n}}$$

- If $\mu_0 \in$ CI, then do not reject H_0
- If $\mu_0 \notin$ CI, then reject H_0 in favor of H_A

Example 3

$CL = 1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

It is claimed that the mean mileage of a certain type of vehicle is 35 miles per gallon (mpg) of gasoline. A random sample of 49 vehicles are collected. The sample mean $\bar{x} = 36.1$ mpg and standard deviation $s = 5.4$ mpg. Use a confidence interval to perform the hypothesis test with $\alpha = 0.05$.

qt(0.975, df = 48) = 2.01

Solution:
$H_0 : \mu = 35$; $H_A : \mu \neq 35$

$$\bar{x} \pm t_{\alpha/2;n-1} \frac{s}{\sqrt{n}} \implies 36.1 \pm 2.01 \frac{5.4}{\sqrt{49}} \implies (34.55, 37.65)$$

Since $\mu_0 = 35$ is inside the interval, we do not reject H_0 at the $\alpha = 0.05$ significance level. We are 95% confident that the population mean miles per gallon is between 34.55 and 37.65.