Lecture 5: Hypotheses Test (I)

Hypothesis: A claim or statement about the value of the unknown population parameter.

when population parameter unknown

A set of hypotheses include two components:

- The null hypothesis; denoted by H_0

Null hypothesis $\quad \boldsymbol{H_0}$: **parameter = value**

- Alternative hypothesis; denoted by H_a

Alt. hypothesis

$$\boldsymbol{H_a}: \textbf{parameter} < \textbf{value}$$

$$\boldsymbol{H_a}: \textbf{parameter} > \textbf{value}$$

$$\boldsymbol{H_a}: \textbf{parameter} \neq \textbf{value}$$

Example 1: A business school claims that the students who complete their MBA training can achieve, on average, **more than** $85000 per year. You want to test this claim.

Claim: mu>85000

H_0: mu<=85000

Null hypothesis: $H_0 : \mu \leq 85000$

Claim: $\mu > 85000$

H_a: mu>85000

Alt. hypothesis: $H_a : \mu > 85000$

Example 2: A customer analyst report that the mean life of a certain type of automobile battery is 74 months. You want to test this claim.

claim: mu=74

H_0: mu=74

$H_0 : \mu = 74$

Claim: $\mu = 74$

H_a: mu != 74 (isn't equal to)

$H_a : \mu \neq 74$

Example 3: A radio station claims that its proportion of the local listening audience is greater than 39%. You want to test this claim.

claim: p>0.39

H_0: p<= 0.39

$H_0 : p \leq 0.39$

Claim: $p > 0.39$

H_a: p>0.39

$H_a : p > 0.39$

Comment: There are some potential mistakes in the setting of hypotheses among some students.

Mistake 1: H_0: $\bar{X} = 10$ vs H_a: $\bar{X} \neq 10$

mistake is the parameter in hypothesis should be population parameter, not sample parameter

Mistake 2: H_0: $\mu = 10$ vs H_a: $\mu \neq 20$

there isn't complement between H_0 and H_a

Tail of the test;
Type of errors.

The tail of test (side of the test)

We use the alternative hypothesis (H_a) to determine the tail of the tests.

- Left tail test: If H_a is " $<$ "
- Right tail test: If H_a is " $>$ "
- Two tail test: If H_a is " $\neq$ "

Example

1. $H_0: \mu = 10 \quad vs \qquad H_a: \mu < 10$

2. $H_0: \mu = 10 \quad vs \qquad H_a: \mu > 10$

3. $H_0: \mu = 20 \quad vs \qquad H_a: \mu \neq 20$

Decision: There are two possible decisions in any hypotheses testing

- Reject the null hypothesis H_0.

- Fail to reject the null hypothesis H_0.

Two types of errors

Example: A person in the court.

H_0: The person is not guilty H_a: guilty

		Actually situation	
		Not Guilty	**Guilty**
Court decision	**Not Guilty**	True	Type II error
	Guilty	Type I error	True

Type I: **Reject** H_0 when it is actually true.

Type II: **Fail** to reject H_0 when it is false.

Significance level; Test statistic; Rejection region

Significance level (α): The maximum probability of the test procedure to have type I error.

The type I and II errors are traded off against each other: for any given sample set, the effort to reduce one type of error generally results in increasing the other type of error.

So what can we do?

First control the more severe mistake (Type I error) and then reduce Type II error as small as possible.

Common significance levels: .10, .05, .01

Test Statistic: A test statistic is a random variable that is calculated from sample data and used in a hypothesis test. You can use test statistics to determine whether to reject the null hypothesis.

Example 1: A person in the court.

H_0: not guilty $\quad$ H_a: guilty

Our test statistic:

Example 2: Let μ be the population mean for a common final exam. One wishes to test

$$H_0: \mu = 70 \quad vs \qquad H_a: \mu < 70$$

What is our test statistic?

One needs to collect a sample data: A sample data of 36 students is selected and sample mean $\bar{x}$ is calculated.

Rejection Region

The rejection region is the interval, that leads to rejection of the null hypothesis.

Example 1: A person in the court.

H_0: not guilty $\quad$ H_a: guilty

Example 2: Let μ be the population mean for a common final exam. One wishes to test

$$H_0:\ \mu = 70 \quad vs \qquad H_a:\ \mu < 70$$

One sample z test for population mean

When to use:

- The sample size is large ($n \geq 30$)

- The target parameter is population mean μ

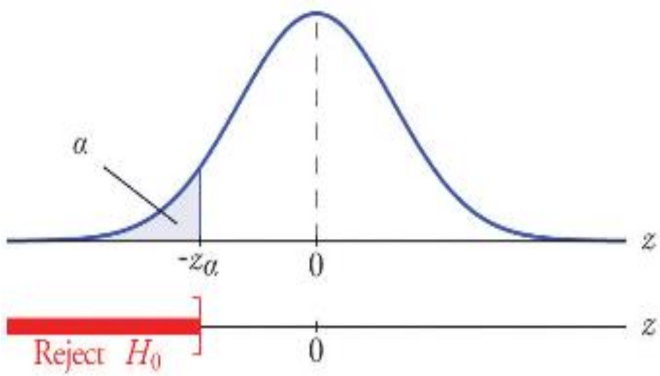
Find the rejection region for z-test

The rejection region for z test will be based on:

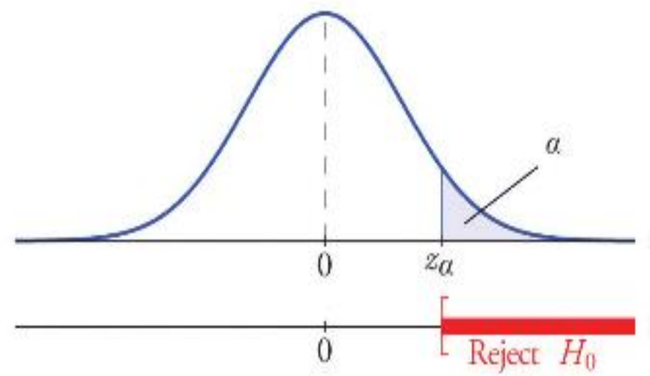
- ❑ The tail of the test

- ❑ The significance level α

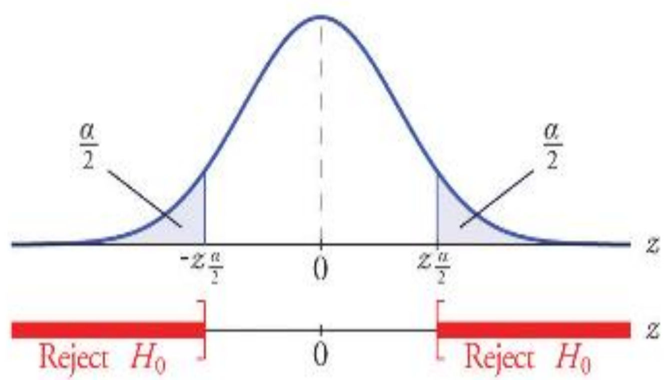
$H_a : \mu < \mu_0$



$H_a : \mu > \mu_0$



$H_a : \mu \neq \mu_0$



a) A left-tailed test with $\alpha = .05$

b) A right-tailed test with $\alpha = .01$

c) A two-tailed test with $\alpha = .05$

Standardized Test Statistic

A **test statistic** measures how far a sample statistic diverges from what we would expect if the null hypothesis H_0 were true, in standardized units.

$$\boxed{z^* = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}}$$

Comment: What is μ_0 ?

Example 1: The CEO of the firm claims that the mean work day of the firm's accountants is **less than** 8.5 hours. A random sample of 36 of the firm's accountants has a mean work day of 8.2 hours with a standard deviation of 0.5 hours. At At $\alpha = .05$, test this claim.

Step 1. Set up the null and alternative hypotheses.

Step 2. The standardized test statistic is

Step 3.The rejection region is

Step 4. Make the conclusion

Example 2: A research study measured the pulse rates of 64 college men and found a mean pulse rate of 70 beats per minute with a standard deviation of 10 beats per minute. Researchers want to know if the mean pulse rate for all college men is **different** from the current standard of 72 beats per minute. At $\alpha = .05$.

Step 1. Set up the null and alternative hypotheses.

Step 2. The standardized test statistic is

Step 3.The rejection region is

Step 4. Make the conclusion