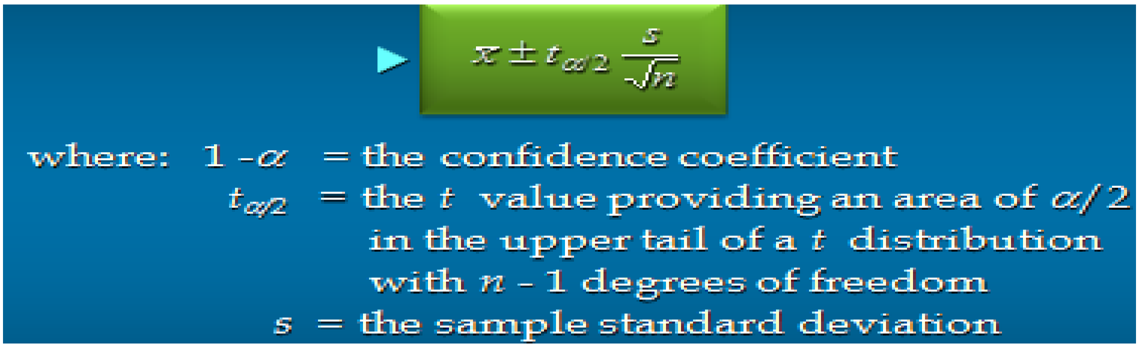
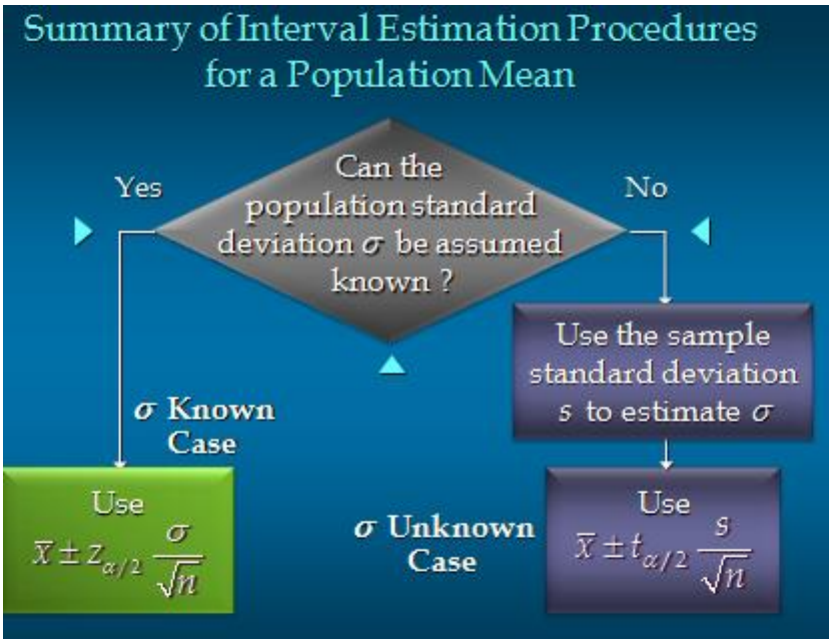
Std is unknown: (t distribution)

- A specific t distribution depends on a parameter known as the degrees of freedom (n-1)
- Refers to the number of independent pieces of information that of into computation of s.
- Distribution with more degrees of freedom has less dispersion.
- As degrees increases, the difference between t distribution and the standard normal probability distribution becomes smaller.



Sample size guidelines are the same.



Lecture – 10:

Hypothesis testing can be used to determine whether a statement about the value of a population parameter should or should not be rejected.

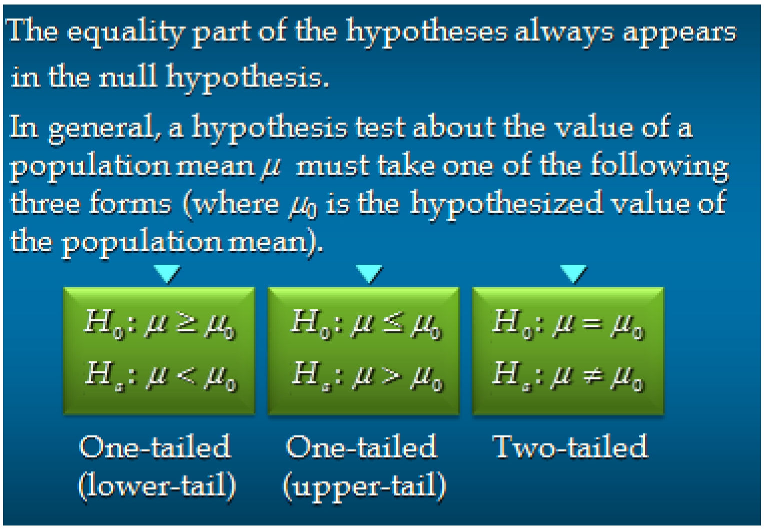
The null hypothesis is a tentative assumption about a population parameter.

Care must be taken to structure the hypotheses appropriately so that the test conclusion provides the information the researcher wants

- In such cases, it is often best to begin with the alternative hypothesis and make it the conclusion that the researcher hopes to support.

Null hypothesis as an assumption to be challenged:

We begin with a belief or assumption that a statement about the value of a population parameter is true. In these situations it is helpful to develop the null hypothesis first.



Type I error:

Type I error is rejecting the null hypothesis when it is true

The probability of making a type I error when the null hypothesis is true as an equality is called the level of significance.

Type II error:

Type II error is accepting null when it is fale.

Very difficult to control the probability for making a type II error

Statisticians avoid the risk of making a Type II error by using “do not reject H_0” and not “accept H_0”.

	Population Condition	
Conclusion	H_0 True ($\mu \leq 12$)	H_0 False ($\mu > 12$)
Accept H_0 (Conclude $\mu \leq 12$)	Correct Decision	Type II Error
Reject H_0 (Conclude $\mu > 12$)	Type I Error	Correct Decision

P value test:

The <u>*p*-value</u> is the probability, computed using the test statistic, that measures the support (or lack of support) provided by the sample for the null hypothesis.

If the *p*-value is less than or equal to the level of significance α, the value of the test statistic is in the rejection region.

Reject H_0 if the *p*-value $\leq \alpha$.

<u>Less than .01</u>

Overwhelming evidence to conclude H_a is true.

<u>Between .01 and .05</u>

Strong evidence to conclude H_a is true.

<u>Between .05 and .10</u>

Weak evidence to conclude H_a is true.

<u>Greater than .10</u>

Insufficient evidence to conclude H_a is true.

- Step 1. Develop the null and alternative hypotheses.
- Step 2. Specify the level of significance α.
- Step 3. Collect the sample data and compute the value of the test statistic.
- Step 4. Use Excel command to compute the *p*-value. (in the exam p-value is given)
- Step 5. Reject H_0 if *p*-value $\leq \alpha$.

Lecture – 11 simple linear regression

- Regression analysis can be used to develop an equation showing how the variables are related.
- The variable being predicted is called the dependent variable – y
- The variables being used to predict the value of the dependent variables is called independent variable – x
- Multiple regression involved more than 2 variables
- The relationship is approximated by a straight line

The equation that describe how y is related to x and an error term is called the regression model.

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where:

β_0 and β_1 are called parameters of the model,

ε is a random variable called the error term.

The simple linear regression equation is:

$$E(y) = \beta_0 + \beta_1 x$$

- Graph of the regression equation is a straight line.
- β_0 is the y intercept of the regression line.
- β_1 is the slope of the regression line.
- $E(y)$ is the expected value of y for a given x value.