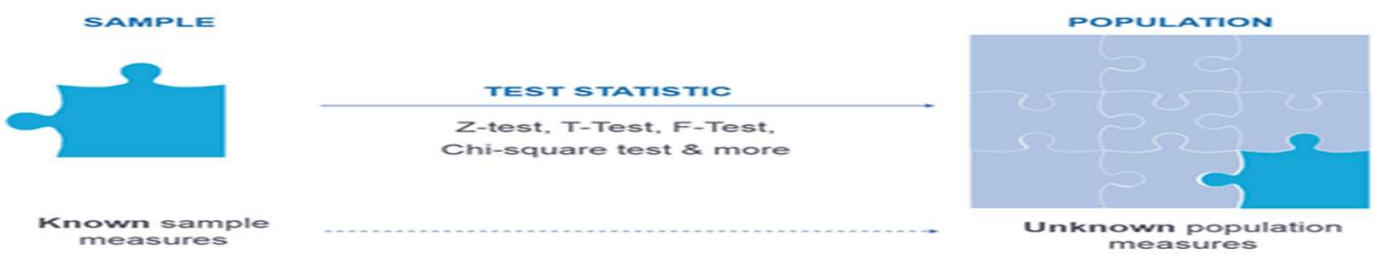
Hypothesis Testing

DR. JASMEET SINGH
ASSISTANT PROFESSOR,
CSED, TIET

What is Hypothesis Testing?

- **Hypothesis testing** is done to confirm our observation about the population using sample data, within the desired error level.
- Through hypothesis testing, we can determine whether we have enough statistical evidence to conclude if the hypothesis about the population is true or not.



Key steps to perform hypothesis test

1. Formulate a Hypothesis: One of the key steps to do this is to formulate the below two hypotheses:
   - **The null hypothesis** represented as H_0 is the initial claim that is based on the prevailing belief about the population.
   - **The alternate hypothesis** represented as H_1 is the challenge to the null hypothesis. It is the claim which we would like to prove as True

2. Determine the significance level: The significance level is the proportion of the sample mean lying in critical regions. It is usually set as 5% or 0.05 which means that there is a 5% chance that we would accept the alternate hypothesis even when our null hypothesis is true

3. Determine the type of test: T-test, Chi Square test, Z test, ANOVA, etc.

4. Calculate the Test Statistic values and the p values

5. Make Decision

Key steps to perform hypothesis test

4. Calculate the Test Statistic values and the p values: A P-value measures the strength of evidence in support of a null hypothesis. If the P-value is less than the significance level, we reject the null hypothesis.

5. Make Decision:

if the **p-value < α**, then we have statistically significant evidence against the null hypothesis, so we reject the null hypothesis and accept the alternate hypothesis

if the p-value > α then we do not have statistically significant evidence against the null hypothesis, so we fail to reject the null hypothesis.

Errors in Decision Making

1) **Type1 Error** – This occurs when the null hypothesis is true but we reject it.The probability of type I error is denoted by alpha (α). Type 1 error is also known as the level of significance of the hypothesis test

2) **Type 2 Error** – This occurs when the null hypothesis is false but we fail to reject it. The probability of type II error is denoted by beta (β)

	Null hypothesis is true	Null Hypothesis is false
we reject the null hypothesis	Type 1 Error	Correct Decision
we fail to reject the null hypothesis	Correct Decision	Type II Error

T-test

- The t test tells you how significant the differences between groups are; In other words it lets you know if those differences (measured in means) could have happened by chance.

- There are **three main types of t-test:**

  1. A One sample t-test tests the mean of a single group against a known mean.
  2. An Independent Samples t-test compares the means for two groups.
  3. A Paired sample t-test compares means from the same group at different times (say, one year apart).

One Sample t-test

- The One Sample *t* Test examines whether the mean of a population is statistically different from a known or hypothesized value. The One Sample *t* Test is a parametric test.
- This test is also known as: Single Sample *t* Test
- The variable used in this test is known as: Test variable
- In a One Sample *t* Test, the test variable's mean is compared against a "test value", which is a known or hypothesized value of the mean in the population. Test values may come from a literature review, a trusted research organization, legal requirements, or industry standards.
- For example: A particular factory's machines are supposed to fill bottles with 150 milliliters of product. A plant manager wants to test a random sample of bottles to ensure that the machines are not under- or over-filling the bottles.

One Sample t-test (Contd....)

Hypotheses

The null hypothesis (H_0) and (two-tailed) alternative hypothesis (H_1) of the one sample T test can be expressed as:

H_0: $\mu = \mu_0$ ("the population mean is equal to the [proposed] population mean")
H_1: $\mu \neq \mu_0$ ("the population mean is not equal to the [proposed] population mean")

where μ is the "true" population mean and μ_0 is the proposed value of the population mean.

One Sample t-test (Contd….)

Test Statistic

The test statistic for a One Sample *t* Test is denoted *t*, which is calculated using the following formula:

$$t = \frac{\overline{x} - \mu_0}{s_{\overline{x}}}$$

where

$$s_{\overline{x}} = \frac{s}{\sqrt{n}}$$

where

μ_0 = The *test value* -- the proposed constant for the population mean
$\bar{x}$ = Sample mean
n = Sample size (i.e., number of observations)
s = Sample standard deviation
$s_{\bar{x}}$ = Estimated standard error of the mean (s/sqrt(n))

The calculated *t* value is then compared to the critical *t* value from the *t* distribution table with degrees of freedom $df = n - 1$ and chosen confidence level. If the calculated *t* value > critical *t* value, then we reject the null hypothesis.

One Sample t-test -Numerical

We have the potato yield from 12 different farms. We know that the standard potato yield for the given variety is $\mu=20$.

x = [21.5, 24.5, 18.5, 17.2, 14.5, 23.2, 22.1, 20.5, 19.4, 18.1, 24.1, 18.5]

Test if the potato yield from these farms is significantly better than the standard yield.

One Sample t-test -Solution

Step 1: Define the Null and Alternate Hypothesis

H0: $\bar{x} = 20$

H1: $\bar{x} > 20$

n = 12. Since this is one sample T test, the degree of freedom = n-1 = 12-1 = 11.

Let's set alpha = 0.05, to meet 95% confidence level.

Step 2: Calculate the Test Statistic (T)
1. Calculate sample mean

$\bar{x}=20.175$
Calculate sample standard deviation

$\sigma=3.0211$

One Sample t-test -Solution

Substitute in the T Statistic formula

T=(20.175–20)/(3.0211/√12)=0.2006

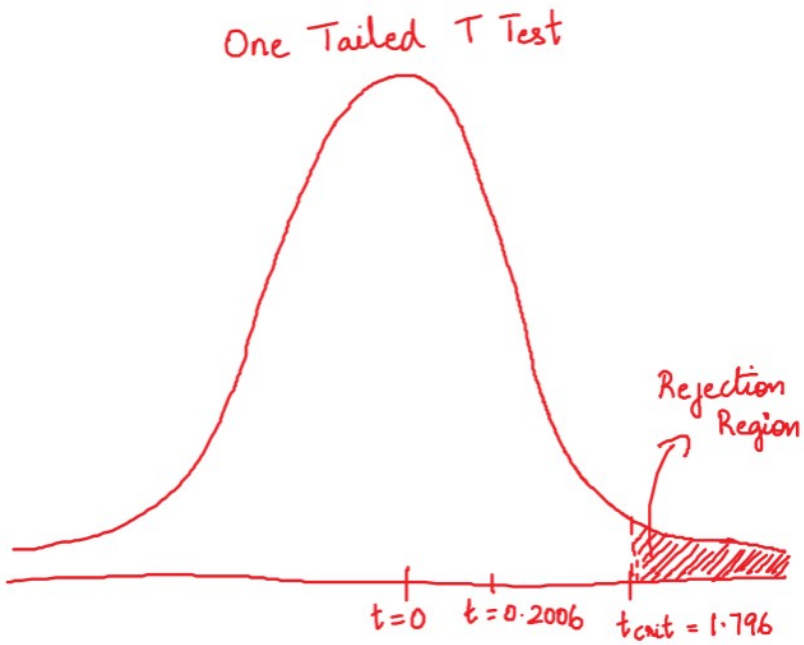
Step 3: Find the T-Critical

Confidence level = 0.95, alpha=0.05. For one tailed test, look under 0.05 column. For d.o.f = 12 – 1 = 11, **T-Critical = 1.796**.

Step 4: Does it fall in rejection region?

Since the computed T Statistic is less than the T-critical, it does not fall in the rejection region.

Clearly, the calculated T statistic does not fall in the rejection region. So, we do not reject the null hypothesis.



Independent Sample t-test

- The Independent Samples *t* Test compares the means of two independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. The Independent Samples *t* Test is a parametric test.

This test is also known as:

- Independent *t* Test
- Independent Two-sample *t* Test
- Student *t* Test
- Two-Sample *t* Test
- Uncorrelated Scores *t* Test
- Unpaired *t* Test

Independent Sample t-test (contd…)

The Independent Samples *t* Test is commonly used to test the following:

- Statistical differences between the means of two groups
- Statistical differences between the means of two interventions
- Statistical differences between the means of two change scores

Independent Sample t-test (contd...)

Hypotheses

The null hypothesis (H_0) and alternative hypothesis (H_1) of the Independent Samples *t* Test can be expressed in two different but equivalent ways:

H_0: $\mu_1 = \mu_2$ ("the two population means are equal")
H_1: $\mu_1 \neq \mu_2$ ("the two population means are not equal")

OR

H_0: $\mu_1 - \mu_2 = 0$ ("the difference between the two population means is equal to 0")
H_1: $\mu_1 - \mu_2 \neq 0$ ("the difference between the two population means is not 0")

where μ_1 and μ_2 are the population means for group 1 and group 2, respectively. Notice that the second set of hypotheses can be derived from the first set by simply subtracting μ_2 from both sides of the equation.

Independent Sample t-test (contd...)

EQUAL VARIANCES NOT ASSUMED

When the two independent samples are assumed to be drawn from populations with unequal variances (i.e., $\sigma_1^2 \neq \sigma_2^2$), the test statistic t is computed as:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where

- $\bar{x}_1$ = Mean of first sample
- $\bar{x}_2$ = Mean of second sample
- n_1 = Sample size (i.e., number of observations) of first sample
- n_2 = Sample size (i.e., number of observations) of second sample
- s_1 = Standard deviation of first sample
- s_2 = Standard deviation of second sample

Independent Sample t-test Numerical

Find the t-test value for the following given two sets of values:

7, 2, 9, 8 and

1, 2, 3, 4?

Independent Sample t-test (Solution)

$\bar{x}1=6.5$, $s1=3.11$
$\bar{x}2=2.5$, $s2=1.29$

Now, apply the formula for t-test value:

$$t = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2})}}$$

$$t = \frac{6.5 - 2.5}{\sqrt{(\frac{3.11^2}{4} + \frac{1.29^2}{4})}}$$

$$= \frac{4}{\sqrt{(\frac{9.3667}{4} + \frac{1.667}{4})}}$$

$$t = 2.38$$

Paired t-test

The Paired Samples *t* Test compares the means of two measurements taken from the same individual, object, or related units. These "paired" measurements can represent things like:

- A measurement taken at two different times (e.g., pre-test and post-test score with an intervention administered between the two time points)
- A measurement taken under two different conditions (e.g., completing a test under a "control" condition and an "experimental" condition)
- Measurements taken from two halves or sides of a subject or experimental unit (e.g., measuring hearing loss in a subject's left and right ears).

The purpose of the test is to determine whether there is statistical evidence that the mean difference between paired observations is significantly different from zero. The Paired Samples *t* Test is a parametric test.

Paired t-test

Hypotheses

The hypotheses can be expressed in two different ways that express the same idea and are mathematically equivalent:

H_0: $\mu_1 = \mu_2$ ("the paired population means are equal")
H_1: $\mu_1 \neq \mu_2$ ("the paired population means are not equal")

OR

H_0: $\mu_1 - \mu_2 = 0$ ("the difference between the paired population means is equal to 0")
H_1: $\mu_1 - \mu_2 \neq 0$ ("the difference between the paired population means is not 0")

where

- μ_1 is the population mean of variable 1, and
- μ_2 is the population mean of variable 2.

Paired t-test

Use the following formula to calculate the t-score:

$$t = \frac{(\Sigma D)/N}{\sqrt{\frac{\Sigma D^2 - \left(\frac{(\Sigma D)^2}{N}\right)}{(N-1)(N)}}}$$

- ΣD: Sum of the differences
- ΣD^2: Sum of the squared differences
- $(\Sigma D)^2$: Sum of the differences , squared.

Paired t-test Example

Calculate a paired t test by hand for the following data:

Subject #	Score 1	Score 2
1	3	20
2	3	13
3	3	13
4	12	20
5	15	29
6	16	32
7	17	23
8	19	20
9	23	25
10	24	15
11	32	30

Paired t-test (Solution)

Subject #	Score 1	Score 2	X-Y	(X-Y)^2
1	3	20	-17	289
2	3	13	-10	100
3	3	13	-10	100
4	12	20	-8	64
5	15	29	-14	196
6	16	32	-16	256
7	17	23	-6	36
8	19	20	-1	1
9	23	25	-2	4
10	24	15	9	81
11	32	30	2	4
		SUM:	**-73**	**1131**

Paired t-test (Solution)

$$t = \frac{(\sum D)/N}{\sqrt{\dfrac{\sum D^2 - \left(\dfrac{(\sum D)^2}{N}\right)}{(N-1)(N)}}}$$

$$t = \frac{-73/11}{\sqrt{\dfrac{1131 - \left(\dfrac{(-73)^2}{11}\right)}{(11-1)(11)}}}$$

$$t = \frac{-73/11}{\sqrt{\dfrac{1131 - \left(\dfrac{5329}{11}\right)}{110}}}$$

$$t = -\ 2.74$$

One-way ANOVA

One-Way ANOVA ("analysis of variance") compares the means of two or more independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different. One-Way ANOVA is a parametric test.

This test is also known as:

- One-Factor ANOVA
- One-Way Analysis of Variance
- Between Subjects ANOVA

The variables used in this test are known as:

- Dependent variable
- Independent variable (also known as the grouping variable, or *factor*)
    - This variable divides cases into two or more mutually exclusive *levels*, or groups

One-way ANOVA (Contd….)

The One-Way ANOVA is commonly used to test the following:

- Statistical differences among the means of two or more groups
- Statistical differences among the means of two or more interventions
- Statistical differences among the means of two or more change scores

One-way ANOVA

Hypotheses

The null and alternative hypotheses of one-way ANOVA can be expressed as:

H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ ("all k population means are equal")

H_1: At least one μ_i different ("at least one of the k population means is not equal to the others")

where

- μ_i is the population mean of the i^{th} group ($i = 1, 2, \ldots, k$)

One-way ANOVA (Contd....)

- The test statistic for a One-Way ANOVA is denoted as *F*. For an independent variable with *k* groups, the *F* statistic evaluates whether the group means are significantly different.

- Because the computation of the F statistic is slightly more involved than computing the paired or independent samples t test statistics, it's extremely common for all of the *F* statistic components to be depicted in a table like the following:

	Sum of Squares	df	Mean Square	F
Treatment	SSR	df_r	MSR	MSR/MSE
Error	SSE	df_e	MSE	
Total	SST	df_T		

One-way ANOVA (Contd....)

where

SSR = the regression sum of squares

SSE = the error sum of squares

SST = the total sum of squares (SST = SSR + SSE)

df_r = the model degrees of freedom (equal to $df_r = k - 1$)

df_e = the error degrees of freedom (equal to $df_e = n - k - 1$)

k = the total number of groups (levels of the independent variable)

n = the total number of valid observations

df_T = the total degrees of freedom (equal to $df_T = df_r + df_e = n - 1$)

MSR = SSR/df_r = the regression mean square

MSE = SSE/df_e = the mean square error

Then the F statistic itself is computed as

$$F = \frac{\text{MSR}}{\text{MSE}}$$