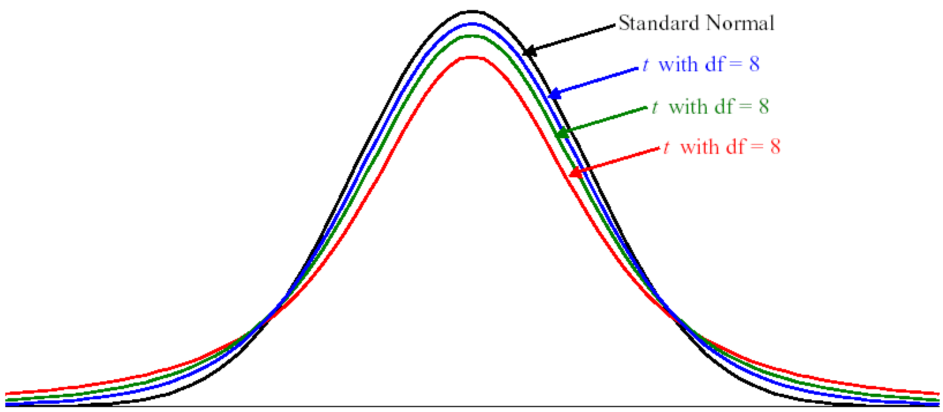
lecture 10

(T) test (student test):-

It's used to study & compare between the mean of sample & mean of population, compare between means of 2 samples, compare between the mean of sample before & after treatment, that if has sample size < 60 <u>can</u> use t test but if equal or below 30 <u>should</u> use t test (= or<30) .

Student's *t*-distribution



The t-distribution has one parameter called the *degree of freedom (df)*, DF=n-1

The t-distribution is similar to the normal distribution:

- Symmetric about 0
- Bell shaped

The main differences between the t-distribution and the normal distribution is in tails (Play around with DF and see the difference of the tails):

- T-distribution has larger tails than the normal
- Larger DF means smaller tails, the larger the DF, the closer to the normal distribution
- Small DF means larger tails

$$t = \frac{x^{-} - \mu}{sd \mid \sqrt{n}}$$

where x^{-} = mean of sample μ = mean of standard population

SD = S.D n = sample size (<30)

<u>Example :-</u>

26 patient after surgery , has standard mean of temp 99Fr^0, S.D =1. standard temp (normal temp of people)= 98F^0

- find if there is statistically deference between temp of patient & normal people

<u>Steps of test</u> :

1. Data :X¯= 99Fr^0, S.D =1 , μ= 98F^0 , n=26
2. H_o : x¯ = μ , H_A: x¯ ≠ μ.
3. Level of significance = 0.05
4. Calculate df = n – 1 = 26-1=25

5. Test statistics: t-test to calculate

6. Assume T distribution

7. Compare t-calculate with t- tabulate :If t-calculate is > t- tabulate, reject the H_o , If t-calculate is < t- tabulate, do not reject H_o.

$\frac{99-98}{1\,|\,\sqrt{26}}$ **= 2.55 (calculated T)**

- Degree of freedom = n-1 = 26-1=25 - Tabulated T = 2.06

- Calculated T (2.55) > Tabulated T = (2.06)

- Reject H_o , p. value → 0.05 → it is statistically significant difference between temp of patients & temp of people after 48 hrs post surgery .

Probability p. value				
degree of freedom	**0.5**	**0.1**	**0.05**	**0.01**
1	**1.000**	**6.31**	**12.71**	**63.66**
5	**0.727**	**2.02**	**2.57**	**4.03**
10	**0.700**	**1.71**	**2.23**	**3.17**
20	**0.687**	**1.71**	**2.06**	**2.84**
25	**0.674**	**1.64**	**2.06**	**2.79**

((T test)) of 2means of 2 samples with equal variance ;

if has sample size < 60 can use t test but if equal or below 30 should use student samples

T test = $(X_1" - X_2") / ((S_{1,2} \times \sqrt{(\frac{1}{n_1} + \frac{1}{n_2})})$

Where $S_{1,2} = \sqrt{\left(\frac{(n_1-1) S^2_1 + (n_2 - 1) S^2_2}{n_1 + n_2 - 2}\right)}$ (find first)

Example; The following data represents weight in Kg for 10 males and 12 females.

Males:

80	75	95	55	60
70	75	72	80	65

Females:

60	70	50	85	45	60
80	65	70	62	77	82

œ= 0.01

Note ; should find Mean & Variance

1. **Data : $Mean_1$=72.7 , N_1= 10 , $Mean_2$=67.17 , N_2 =12**

   $Variance_1$ =128.46 $Variance_2$=157.78

2. **H_o: $mean_1 = mean_2$, H_A: $mean_1 \neq mean_2$.**

3. **Level of significance (alpha) = 0.01**

4. **Calculate df = $n_1+n_2 - 2 = 20$**

5. **Assume T distribution**

6. **Test statistics: t-test to calculate.**

7. **Compare t-calculate with t-tabulate :If t-calculate is > t- tabulate, reject the H_o , If t-calculate is < t- tabulate, do not reject Ho.**

$$T\ test = \frac{X1'' - X2''}{S_{1,2} \times \sqrt{\left(\frac{1}{(n_1)} + \frac{1}{(n_2)}\right)}}$$

$$S_{1,2} = \sqrt{\left(\frac{(n_1-1)S^2_1 + (n_2-1)S^2_2}{n_1 + n_2 - 2}\right)}$$

$$S_{1,2} = \sqrt{\left(\frac{(10-1)\times 128.46 + (12-1)\times 157.78}{10 + 12 - 2}\right)}$$

$$S_{1,2} = \sqrt{\left(\frac{1156.14 + 1735.58}{20}\right)} = \sqrt{(579.44 / 20)} = \sqrt{144.586} = 12.024$$

T test = 72.7 - 67.17 /(12.024 x √ (0.1 + 0.083)) = 5.53 / 12.024 x √ (0.183)

= 5.53 / (12.024 x 0.428) = 5.53 / 5.146

t = 1.075

- The tabulated t, for alpha 0.01 is 2.84

- the calculated t (1.075) < tabulated t (2.84) , p>0.01

Then accept Ho and conclude that there is no significant difference between the 2 means, this difference may be due to chance.

3 . <u>Compare after & before treatment of one sample with 2 occasions)</u>:

$$Td = \bar{d}_{mean} / S\bar{d} / \sqrt{(n)}$$

$$S\bar{d} = \sqrt{\left(\frac{\sum d^2 - (\sum d)^2 / (n)}{n - 1}\right)}$$

df = n – 1 , $\bar{d}$ = mean , $S\bar{d}$ = SD of deference between after & before

<u>Steps of test</u> :

1. Data : $\bar{d}_{mean}$ = mean , d^2 , $(\sum d)^2$

2. testing hypothesis :

H_o: there is no significant difference between readings before and after treatment
H_A: there is significant difference between readings before and after treatment

3. **Level of significance**

4. **Calculate df = n – 1** 5. **Assume T distribution**

6. **Test statistics: $Td = \bar{d}_{mean} / S\bar{d} / \sqrt{(n)}$**

7. **Conclusion : compare t-calculate with t-tabulate :If t-calculate is > t- tabulated reject the Ho , if t-calculate is < t- tabulate, do not reject Ho.**

$Td = \bar{d}_{mean} / S\bar{d} / \sqrt{(n)}$ Where $S\bar{d} = \sqrt{\left(\frac{\sum d^2 - (\sum d)^2 / (n)}{n-1} \right)}$

Example: Systolic Blood pressure of 8 patients, before & after treatment

Before	After	d	d^2
180	140	40	1600
200	145	55	3025
230	150	80	6400
240	155	85	7225
170	120	50	2500
190	130	60	3600
200	140	60	3600
165	130	35	1225
Mean= $\bar{d}_{mean}$ = 465 / 8 = 58.125		$\sum d$ =465	29175 $\sum d^2$

1. **Data : $\bar{d}_{mean}$ = 58.125 , d^2 = 29175 , $(\sum d)^2 = (465)^2$, n = 8**

2. **testing hypothesis :**

H_o: there is no significant difference between readings before and after treatment
H_A: there is significant difference between readings before and after treatment

3. **Level of significance = 0.05**

4. **Calculate df = n – 1**

5. **Assume T distribution**

6. Test statistics: Td = d¯ mean / Sd¯/√ (n)

7. Conclusion : compare t-calculate with t-tabulated :If t-calculated is > t- tabulated then reject the Ho , if t-calculated is < t- tabulated, do not reject Ho.

Td = d¯ mean/ Sd¯/√ (n)

$$Sd^{-} = \sqrt{\frac{\sum d^2 - (\sum d)^2/(n)}{n-1}}$$

$$Sd^{-} = \sqrt{(29175 - ((465)^2/8)/7)} = \sqrt{(29175 - 27028.125/7)}$$

$$= \sqrt{(2146.875/7)} = \sqrt{306.606}$$

Sd¯=17.510

Td = d¯ mean / Sd¯/√ (n)

$$= 58.125 / 17.510 / \sqrt{8} = 58.125 / 17.510 / 2.83$$

$$= 58.125 / 6.19 = 9.39$$

-Tabulated t (df 7), with level of significance 0.05 = 2.365 (from table)

- Calculated t > Tabulated t P value < 0.05

-We reject H_o and conclude that there is significant difference between BP readings before and after treatment at level P <0.05.

Which T-test to use : للاطلاع

Type of T-test	**Assumption**	**T**	**DF**
one-sample t test	Usually used to compare the mean of a sample to a know population	$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$	**n-1**
Two-sample	The **variance** in the two groups are **extremely different**. e.g. the two samples are of very different sizes unequal variance (heteroscedastic t-test	$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_1} + \frac{S_Y^2}{n_2}}}$	$df' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

Two-sample assuming	Two samples are referred to as independent if the observations in one sample are not in any way related to the observations in the other equal variance (homoscedastic t-test)	$$T = \frac{\overline{X} - \overline{Y}}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$	n1+n2-2
Compare two paired groups	used to compare means on the same or related subject over time or in differing circumstances; subjects are tested in a before-after situation	$$T = \frac{\overline{X} - \overline{Y}}{S_d / \sqrt{n}}$$	n-1

Confidence intervals for (T) test ;

To estimate the population parameter: at 95% ,99% , 90% C. Level

1. **one sample of one mean :**

   X” ± (t df n-1) × (SD/(√(n)

2. **Two samples of two means :**

   (X_1”- X_2”) ± (t df n_1+ n_2 –2) × Sp × √ (1/n_1 + 1/n_2)

where Sp = $S_{1,2}$ = √ ((n_1 - 1) S^2_1 + (n_2 - 1) S^2_2 / n_1 + n_2 – 2)

3.(t) after & before :

d¯ mean ± tdf n-1 × (Sd¯/(√(n)