***T- test***

measures the differences between **two group** means (not frequencies and percentages as in non parametric stat).

* main assumption of using t test is that the level of measurement should be **continuous level** as we are dealing with means the second assumption is that we should have **normal distribution** (if it’s skewed we can’t use parametric stat and we should use the alternatives such as the non parametric statistics).

T test mainly comparing group means not individuals means

For example two groups of student we compare between means of each group by using the same teacher and we have to see which one is better

Another example two groups of diabetic patient type 2

We compare between hba1c for each group and measure for example the effect of oral hypoglycemic and another regular medicine

So we will see if the drug effective or not

If we have any violation in this assumption we go for another test

What is the probability of getting a difference this large by chance only

The most important thing to concern about the dependent value should be continuous either ratio or interval

Non continuous either mean mode median standard deviation about variability nominal only mode

the most important thing to think about is normal distribution of continuous dependent variable

Normal distribution: the mean .mode, median at the same level zero level skweness less than 2 SD

So if dependent variable for example hba1c normally distribution we can use t test

If it not normally distribution we have another options

-non parametric statistic:

If we have hba1c mean 9 mg/dcl we categorize it in two groups nominal distribution 1-above mean 2- under mean and because we have two groups nominal distribution we use chi-square.

-data transformation most difficult: we measure skweness and z score if skweness was + Subtract

 from mean and if it negative addition to the mean

* We choose group randomly so the most of subjects will be in the same circumstances
* Assumptions

- Two mutually exclusive groups

- Normally distributed dependent variable

- Homogeneity of variance

-Intervention: new device ,new material ,new way of treatment.

Control means that you should have experimental group:

Blocking technique: we need all subject female for example in breast cancer

Matching technique: match between health condition of group member

 So these 3 conditions if found it will give us experiment characteristic.

So one of them is very important in performing the T-test is that the two groups ( experimental and control ) we utilize, there should be sore of compatibility between them.

When we do the t-test, the SPSS as a software there is an option that we can do examine or test, the SPSS do it already by itself on data **homogeneity of various** or it called [**homoscedasticity**](https://en.wikipedia.org/wiki/Homoscedasticity).

**What is the meaning of homogeneity of various ?!**

This mean that I do not want to see significant variations between the characteristic of this group compared the experimental and control.

If the homogeneity is significant, this mean that utilizing t-test is not effective ( using of t-test is not valid in this way ), so you should go and do reduction to these variation.

**There is two types of t-test:**

1 ) one sample of t-test:

- it compare between the product and itself or with its criteria.

- it is used in trading.

- for example if I am factory manager and I want all the product boxes exhibit special characteristic so I found the criteria difference and I do one sample and I do the test on it.

- other example the researcher may want to test wither average IQ score for four group IQ for students according to 100.

- other example compare between weight of the pocket, that the weight should be around 1.3 at level of 95% confidence interval and the alpha=0.05

The doctor explain what is the meaning of confident interval that you are confident about this level, so in 95% I am confident that at least 95% of the subject are examine and this is the minimum level of examinant the sample, because some of the research especially the RCT it is not acceptable that alpha is 0.05, it should be lesser than this, it should be 0.01 or 0.001.

Some research is very strict such in math or physics.

So if the confident interval was 94%, the alpha will be 0.06 and this is not acceptable, so the smaller the alpha was the results will be more strict, but in this case you need large sample and you need more characteristic.

2 ) two groups of t-test ( independent test ):

- most common used.

- two independent groups.

- experimental and control.

From the science of statistics we have the science of **power analysis** which include four elements:

1 ) alpha significant level.

2 ) power itself ( power of the test ).

3 ) effect size, it is the degree of which H0 is false, which mean the value of the relationship actually between dependent and independent usually it is measure by small, medium or large, the smaller the smaller the effect size, the bounded well be better, but this is very expensive.

4 ) the sample size.

These four elements always bonded together in the measurement, so if we know three of them we can found the fourth from them, so we use specific formula for this, most common we use it to know how much formula we need.

Also after we finish the research, we can use the sample size, effect size and alpha to find the power.

So we always reach compromise stage ( intermediate result between them ).

The people who do the research in general from the last century, they give us general role and general guidance that the power should be 80%, alpha 0.05 and effect size is medium, in this case your result will be acceptable.

If you want to make it better, you should improve these indicators, the effect size improve it from medium to small, the power should be improve to more than 80% ( 90% - 99% ) and increase sample size.

There is no study in the world you are sure 100% that your result are true.

There is specific software for them it called je star power, this is science by its own called power analysis.

So after each test we should use the power analysis.

**Now go back to t-test:**

Example of one sample t-test ( the doctor tell us that its not required from us, it is just for our knowledge ).

**The test that is required from us is the independent sample t-test:**

-it mean two independent groups which compare the means between two groups of cases.

- ideally for this test subject should be randomly selected, so that any differences in responses view of the treatment or lack of treatment on control group (placebo) or student treatment on experimental group and not to the other factors such as the age ( unusual differences ), Differences in average income may be influence by factors such as education and not by gender alone.

So if you are compare between different of income in two groups, you should see if there is effect or increase or specific increment for employers in specific company.

Example: they do study on doctors university about education level and the salary, her the salary is independent variables, so as the education level increase, the person look to the money decreased.

Or the gender, the women is not looking to the salary as the man.

Other example; patient with high blood pressure are totally assigned to the placebo group and treatment group, so her the dependent variables her is the blood pressure.

Now we will try if the placebo subject receive an inactive pills, the placebo are three types sugar or starch or capsule that are resemble to the other capsule in the shape but it contain starch and sugar instead of the drug or it may be other drug less effective than the drug we are test it or nothing.

Which of these is ethical ?? the use of drug which is less effective.

So the control group take native pill placebo, and the treatment subject receive the new drug to reduce the blood pressure that is expected to reduce the blood pressure.

After treatment of the subject for two months and the two samples t-test is used to compare the average of blood pressure for the placebo group under the treatment and the other group and each patient is measures once and belongs to one group.

So in this example the result was the average mean of the reduction of blood pressure among the experimental group is highly different from control group we utilize the t-test.

Other example is the student perception with the differences by gender, showed that females students have higher score of perfect perception.

So the researcher is reporting the **value of t** and usually it is positive or negative also there is the degree of freedom (d.f.) which is n-1=603 in this example, its means that the whole population of sample is 604 and the p value = 0.00 or less than 0.001

If the result of t-test was negative, this mean that the group number 2 has the higher result.

And if the result of t-test was positive, this mean that the group number 1 has the higher result.

So if I enter the experimental group first then the control group and the t result was in negative this mean that the control group has the higher result ( so this significant is render to the control group ) and vice versa.

So they found that the reduction of blood pressure is better in the control group, so there is difference between the two groups, so there is significant differences between the groups and we reject the null hypothesis and accept the alternative hypothesis.

If the experimental group is the second group then the new drug is effective in the reduction of blood pressure, so we reject the null hypothesis and accept the alternative hypothesis ( research hypothesis ) and the superiority go to the experimental group.

We should always look to the p value first then to the t value, if there is significant differences between the groups then we go to the t value to know to whom this difference, but if there is no significant difference so there is no need to do t-test and there is no need to know the result of t-test.

The p value is never 0.00 it will be less than 0.001, and if its less than 0.001 there is significant difference and you should look to the t value.

So the value of t-test is not important but it is important to know the direction ( positive or negative ).

لازم تعرف مين المجموعة رقم 1 ومين المجموعة رقم 2 على حسب إدخالهم للاختبار ..

The next lecture will be about the example of t-test and ANOVA and MANOVA.Sorry for any mistake, wish you the best.. ☺Alaa Ali