Lecture 13

In the previous lecture, it was illustrated that the difference between the parametric and the non-parametric is in whether the parametric criteria or assumptions has been met or not, and I**f** there is **any violation in any** of the parametric assumptions, **then** we should use the non-parametric tests.

The **criteria (assumptions)** to using ***parametric*** tests are about assuring the sample being:

1. **Normally distributed**.
2. **Not skewed**.
3. **Adequate sample size in both groups.** (*additional "you can ignore it" note*: it needs adequate sample size in order to obtain the adequate degrees of freedom. *Degrees of freedom (df): The number of values that are free to vary if the sum of the total is given "will not significantly change the mean if they change within this range value.") …

-***Parametric*** tests: work with normally distributed scale data.
-***Non-parametric*** tests: used with not normally distributed scale data and with ordinal data.

This lecture will include two common examples on statistical tests, ***t-test*** and what corresponds to it in the non-parametric type of tests, ***Mann-Whitney test***.

t-test

t-test is a parametric test, and is the most commonly used test in research. Always it compares between the **means** of TWO groups; young vs. elderly, smokers vs. non-smokers, or females vs. males. And there are different versions of t-test for different designs. While the non-parametric tests compare between the **median** of the TWO groups. Nonetheless, they both **don't** use to compare between each individual case within the 2 groups, but instead they use one of the central tendency measures, parametric → mean, and non-parametric → median.

Types of t-test:

1. **Un-paired** → Dependent (related) samples "before & after":

–compare the means of two conditions in which the same (or closely matched) participants participated. Example: The score of some teenagers before the educational program and after it. Before the educational program the mean of the scores was 69.2 and after the program it became 78.2, so the difference is 9. Then to determine whether it's significant or not the SPSS will calculate the p-value, which will determine whether to refute the H_0 or not.

2. **paired** → Independent (unrelated) samples:

–compare the means of two groups of participants. (E.g. the difference in weight loss between a group of males and a one of females along a period of 3 months).

- t- test must have 2 variables, **Independent variable (IV)** & **Dependent variable (DV):**

1. The **IV** must have two related or unrelated groups and only two, an intervention group and a control group.

2. The **DV (the outcome)** must be a continuous variable, interval or ratio, though it could be an *artificial dichotomy* (distributing the values into two ranges, then referring to all values within one range as being the same within one category, young "<65" vs. elderly ">65").

Research is built on statistics, and in the comparative research design we should prove that there is significant difference statistically. Using t-test we're targeting significant differences between the samples' means, assuring whether it's a false (due to chance or error) or a real difference. Remember we're looking for differences not associations as opposed to the correlational type of statistics.

Generally:

1. ***Null "statistical" hypothesis (H_0)***: states that the samples come from the same population. No significant difference between the 2 groups.
2. ***Research "alternative" hypothesis (H_A)***: states that the samples come from different populations. There is a significant difference between the 2 groups.

SPSS (Statistical Package for the Social Sciences)

Most commonly used statistical program around the world. It is easy and very efficient for analyzing, organizing, and presenting data. However, what it only needs, is to be frequently practiced in order to master it.

In the SPSS, there're several windows: the data view, the variable view, and the output view.

The data view displays your actual data and any new variables you have created. In SPSS, columns are used for variables, while rows are used for cases (also called records).

The variable view found at the bottom of the data window. It contains the definitions of each variable in your data set, including its name, type, label, size, alignment, and other information.