

Level : Master I (LS)

Module : Methodology in Social Sciences

# *Data Analysis*

Prof. Rahmouna ZIDANE

Department of English

University of TLEMCEN

# *Data Analysis*

- ◆ Data analysis is an important step in reporting the research findings. It may be qualitative or quantitative depending on the type of the collected data.

# Quantitative *Data Analysis*

- ◆ Quantitative data analysis relies on a set of statistical techniques that are employed for the interpretation of the collected information. It involves the use of descriptive and inferential statistics.

# Descriptive Statistics

- ◆ Descriptive statistics aim at describing and summarizing the gathered data. They include frequencies, percentages and measures of central tendency as well as measures of variability.

## Descriptive Statistics

- ◆ Frequencies refer to enumerative data or numbers which can be transformed into percentages. Frequencies and percentages may be provided within a text or they may be represented in tables called frequency tables.
- ◆ Percentages may be demonstrated through graphic representation which involves the use of bar-graphs or pie-charts.

# Measures of Central Tendency

- ◆ Measures of central tendency are used for the description of the central point of a set of data. These statistics are : the mean, the median and the mode.

## Measures of Central Tendency

- ◆ **The mean** is the average score of a group of individuals. Its formula is as follows:

$$\bar{x} = \frac{\sum x}{n}$$

$\sum x$  = the total of scores

$n$  = the number of scores

For instance, in the following list of scores: 2, 3, 5, 6, 8, 10, 12, 14, 15, 16, the mean is (9.1).

## Measures of Central Tendency

- ◆ **The median** shows the middle value of a group of scores. For an odd number, the median is the middle score; For example, in a list of scores including the following values: 5, 6, 8, 11, 12, 14, 15, the median is 11.
- ◆ For an even number, the median is the average of the two scores which are in the middle; for example, in a list of scores including the following values: 2, 3, 5, 6, 8, 10, 12, 14, 15, 16, the median is 9  $[(8+10)/2]$ .



## Measures of Central Tendency

- ◆ **The mode** is the number that frequently occurs in a distribution of scores. For instance, in the following list of values: 6, 8, 10, 9, 12, 8, 14, 8, 11, 8, the mode is 8.

# Measures of Variability

- ◆ Measures of variability are statistical procedures that are employed to identify the dispersion of data around the mean; they denote the spread of a set of values . Measures of variability include the range and the standard deviation.

## Measures of Variability

- ◆ **The range** refers to the difference between the highest and lowest value in a dataset. Its formula is as follows:  
Range = (largest value - smallest value)

For example, in a list of scores including the following values: 2, 3, 5, 6, 8, 10, 12, 14, 15, 16, the range is 14 (16-2).

## Measures of Variability

- ◆ **The standard deviation** is a measure of variability that determines the dispersion of data; it helps to show if the values are close to or far from the mean.

## Measures of Variability

The formula of the standard deviation is as follows:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

$s$  = standard deviation

$x$  = each score

$\bar{x}$  = the mean

$n$  = number of scores

# Inferential Statistics

- ◆ Inferential statistics refer to the statistical procedures that are employed to draw conclusions about the features of the target population depending on the data collected from the selected sample. They aim at making predictions.
- ◆ Inferences can be made from quantitative data through the use of inferential statistics such as the chi-square test.

## Inferential Statistics

**The chi-square test** is a very useful technique that helps to determine the difference between the observed and expected data. Its formula is as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$O$  = observed frequencies

$E$  = expected frequencies

# Quantitative *Data Analysis*

- ◆ Quantitative data analysis represents an important step that enables the researcher to summarize the research findings in an accurate and objective way. It provides numerical data that can lead to make inferences related to the generalization of the obtained results.



# References

- ◆ Cohen, Louis, Manion, Lawrence, and Morrison, Keith. (2007). *Research Methods in Education*. New York: Routledge.
- ◆ Fleming, Michael C. and Nellis, Joseph G. (2000). *Principles of Applied Statistics*. UK: Thomson Learning.