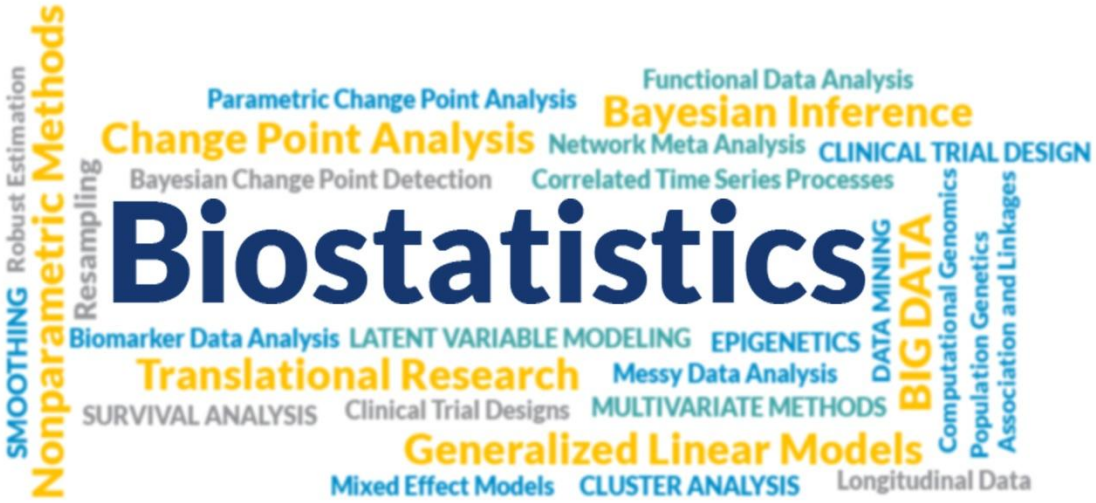


◉Sheet ○Slides

Number

...

Done by:

Mariam Hassouneh

corrected by:

Rasha Al Mousa

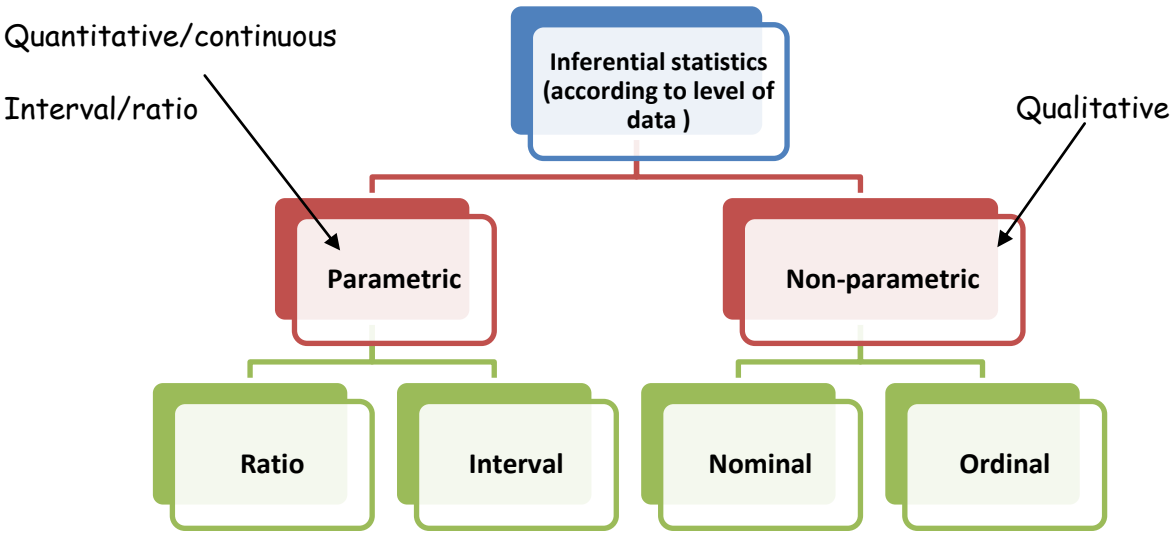
Doctor

Mhmoud Al hussami

Today's lecture will be the **last lecture** included in the Midterm exam.

Slide 6

The first 10 slides are covered (and copied as they are).



1. **Parametric techniques:**

These techniques are statistical "have some rules to put", so any violation " انتهاك " for any of the rules will assure their invalidity to be used, and so non-Parametric will be used instead!

- What are these rules?

1. <u>The observations must be independent</u>: for example, a score for a midterm exam is independent than a score on a final exam. Same goes for the GPA at the end. You cannot assign the same score on the midterm for the final and GPA. Each observation must be sorted on its own and independent than others.

2. <u>Dependent variable</u> (the outcome) should be <u>continuous</u> (Intervals/Ratios).

3. <u>The observations (the data coming from dependent variable) must be drawn from normally distributed populations</u>. This applies if we were working with an arithmetic mean AND geometric mean. Even if the data was not normally distributed and we converted it to get a geometric mean, we still can use parametric techniques.

*Parametric techniques are not to be used with data that has outliers. Data has to be trimmed.
** Geometric mean isn't included in the exam, but just be familiar of this point.

4.<u>These populations must have the same variance</u> (*recall that the variance is the standard deviation squared, and the standard deviation describes how deviated the data are about the mean*): In order for the sample to be representative of the population, the variance has to homogenous or similar throughout different groups of the population. This way, if the data follow a certain pattern in the sample, they should follow that same pattern in the population as well.

Student's question: same with the population or also with other samples?
Answer: With the population. Between samples we find the standard error of the mean (St. dev / root of the sample size). If it's a big value, then there is no significance and the group is heterogeneous. Although the population should be homogenous referring to the known rule as **"Population traits ought to be normally distributed among the population"**

5. <u>The groups should be randomly drawn from normally distributed and independent population</u>

- A randomly drawn group is created using true experimental techniques, not Quasi Experiments.
- In a Quasi, there are **no** randomly assigned groups. Each sample has no variation. It might consist of a group only composed of doctors or a group of pharmacists. There is no overlap in a quasi.
    - Random selection in a quasi can happen by assigning numbers, all who carries an odd number will be grouped together, same for those with even.
- The independent variable must be categorical(2 or more elements should make it up). The distribution of these two or more variables should be normal, just like in the population(matching). If there's data with a GPA=4 in the population, then it should exist in the sample.

To remember: when doing a test, what determines what you'll be using is considering if:

1. Manipulation is allowed or not.

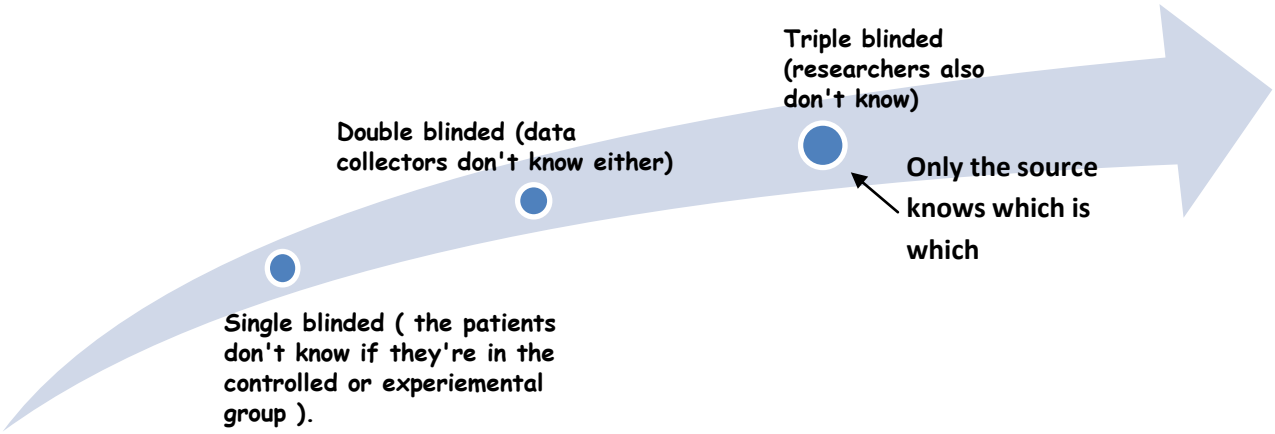
2. Randomization is necessary or not.

3. You had to control the sample and group people to those who have a condition vs. those who don't.

If 1, 2 and 3 are to be used, it's a true experimental (which the doctor is talking about)

If 1 and 3 , Quasi is used.

Blindness. We have 4 levels:



2.38min – 12.00min

- Why to use Parametric?

1. They are more powerful (results are excellent and results I want to apply on reference population is realistic), more flexible, and easier to calculate than nonparametric techniques.

2. They don't only allow the researcher to study the effect of many independent variables on the dependent variable, but they also make possible the study of the interaction between the independent variables.

** We took previously that statistics when testing variables are of three types:

1. Univariate statistics: Studied by descriptive only because it's a single variable.
2. Bivariate statistics: Studied by descriptive and inferential but with considering no interaction (the independent variable is the only effect to be studied).
    a. If we had dependent and independent variable, we'll study the effect of the independent (tobacco use) on the dependent (lung cancer).
3. Multivariate statistics: Studies interactions between variables together (advanced).
    a. For example, If we had a dependent variable with **a lot of independent variables** affecting it, we'll also consider the interactions between the independent variables and how they **together** affect the outcome. Like heart diseases, we'll find types of food, smoking cigarettes, behaviors and many others affect how much the person is prone to heart diseases. They are studied by Manova (is a procedure for comparing multivariate sample means, used in case we want to study the effect), and multiple regression (used in the case of prediction).

2. Non- parametric techniques:

- Nonparametric methods are often the only way to analyze nominal (percentages) or ordinal data (Median is studied) and draw statistical conclusions.
- For nominal analysis-> Chi-square.
  For ordinal analysis→ Mann Whitney U-test and Kruskal-Wallis
- No need for a normal distribution(distribution-free methods)
  - Since in skewed distributions the mean is affected, we use the median here(ordinal data)
- We can use them in small sample sizes(results aren't generalized unlike parametric that require a large sample size)
- They have no assumptions "rules to put" like parametric.

- In nominal data, nothing but associations are to be studied, causality isn't studied from methods used with nominal data.

Chi – square is used in percentages , suppose we have 2 different groups (males , females) , and we would like to see their opinion in a thing, such as their agreement on junk food whether they like it or not, notice that there is **absolutely nothing** to find a cause or effect from, just percentages and sometimes modes of samples are calculated. A sample of 20 people in Chi-square would be enough.

When using SPSS you want to measure anxiety level, Mann Whitney or Kruskal Wallis (In the table , last page) are used. You'll have to order as following :
1: mild 2: moderate 3: severe here the software can recognize data arrangement.

If you tried to place them haphazardly as following:
1: moderate 2: severe 3: mild the software will not obey, will **rearrange** according to what it's programmed to (as the 1st order), and will bring wrong results against what you're expecting!

- Thus, with ordinal data, make sure to order the data or else the technique will not figure them out and results will be wrong!

* 12.00min – 20.20min*

- There is at least one nonparametric test equivalent to a parametric test

A VERY IMPORTANT POINT: For the sake of the exam, you'll have to understand the following notes because questions won't be direct (he won't ask to get mode or median), it'll be applied on real life use of these techniques.

There are several types of techniques used in analysis:

1. Tests of differences between groups (independent technique):

Be sure that when you're testing differences between groups, you're using the independent techniques (to be seen in the table)

2. Tests of differences between variables (dependent technique):

When you're testing between variables within a group, then you're using dependent techniques.

3. Tests of relationships between variables:

If you wanted to study relationships (associations), then here for every participant you'll have to match dependent with independent (because there's no such a mean). HOW THEN? 120 students, the first got a grade of 60%, the last got a grade of 100%, then we randomized the grades, T test and ANOVA will not be affected because they depend on the mean, and even if grades where random **,the mean will never be affected**!

So even also if we compared the mean of the group with other groups, nothing will happen. Now if we wanted to test associations of variables for the same person, and his mark was placed randomly (not the true one), this would be wrong.

<u>Conclusion</u>

1. Parametric Statistics are statistical techniques based on assumptions about the population from which the sample data are collected.

    A. Assumption that data being analyzed are randomly selected from a normally distributed population.

    B. Requires quantitative measurements that yield interval or ratio level data.

2. Nonparametric Statistics are based on fewer assumptions about the population and the parameters.

    A. Sometimes called "distribution-free" statistics.

    B. A variety of nonparametric statistics are available for use with nominal or ordinal data.

* 20.20min -28.20min*

- A lot of questions will come on the table below, it is **all required** as names but details for some methods are to be known later on. I'll try to simplify as much as possible. You can find some notes here:

1* This table classifies tests according to levels of Data (vertically) and number of groups (horizontally).

2* For each horizontal classification for number of groups, there is dependent and independent part, taking into account if you're comparing between :

1. Groups, like between females and males , 2 different groups with no overlap (independent techniques , page 6)
2. Variables within the same group , females or males level of anxiety (2 variables here), you'll measure the level before and after the exam (Dependent, page 6)

3* For nominal data, Chi-square is the main analytical type. For dependent we use Goodness-of-fit, for independent we use independence Chi-square.

4* For ordinal data, Mann Whitney is used when comparing 2 medians of 2 separate groups, and Wilcoxin is used when comparing 2 medians of the same group before and after(dependent).
if they were 3 or more groups, independent → kruskal-wallis
if they were 3 or more groups, dependent →

5* When testing Correlations of continuous with ordinal, we use Point-biserial. Other than that and the mentioned in the table, we use T test or chi- square (it studies association).

6* Factorial (2-way) ANOVA is used when studying interactions of different independent variables within groups. Here we have more than an independent variable (refer to page 4, multivariate)

Summary Table of Statistical Tests

Level of Measurement	Sample Characteristics					Correlation (Associations)
	1 Sample	2 Sample		K Sample (i.e., >2)		
		Independent	Dependent	Independent	Dependent	
		Chi-square	(Goodness of fit)			
Categorical or Nominal	X^2 or bi-nomial	X^2 The main Method used with nominal	Macnarmar's X^2 (Also a Chi – square)	X^2	Cochran's Q	
Rank or Ordinal	There is no technique	Mann Whitney U	Wilcoxin Matched Pairs Signed Ranks	Kruskal Wallis H	Friendman's ANOVA	Spearman's rho
Parametric (Interval & Ratio)	z test or t test (Standard tests)	t test between groups (The independent T-test)	t test within groups (The dependent T-test)	1 way ANOVA between groups	1 way ANOVA (within or repeated measure)	Pearson's r
		Factorial (2 way) ANOVA				

With it we mainly depend on percentages

Continuous

Different from 1-way, 2-ways that are used with continuous data

Within the group

28.20min -35.45min --> Here ends the mid exam material ^.^

Please forgive me if anything wasn't clear. Some questions answers and words I just didn't manage to hear nor predict! I preferred to only write things I hope I heard right and to explain what I'm sure of. Wishing you now to take a deeep breath, and I guess I'll imagine as well a wide smile :D . Kindly pray for your sister, Blessed Cure as always ♥