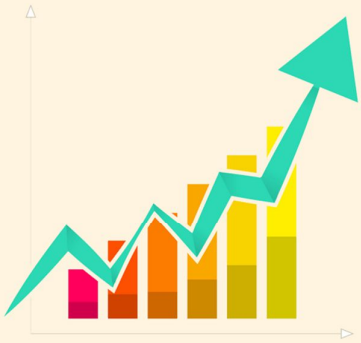
Biostatistics

Doctor 2017 | Medicine | JU



Number >>	9
Doctor	Hamza Alduradi
Done By	Nael jadallah
Corrected By	Rawan Almujaibel

This sheet was written from section 3 record

Last time we finished unit 3 which is about probabilities

In previous lectures, we have talked about the Z table and how to use it (it was the simplest example of transition from the descriptive statistics stage to the inferential statistics stage and how we can complete the scenario of biostatistics)

Steps:

Start with Calculating mean and standard deviation → extract the Z score by inferential statistics calculation → go to Z table and use the number we calculated to find the probability → finally we end up with an answer to our research question .

But research question that Z table answer, (the study was the probability of the women who is their height is 140 cm or less) it is very simple because it has 1 variable only (ex. the height of the women) and this is uncommon because it doesn't have any comparison, hypothesis, and doesn't have any challenge. In fact the most common type of research have at least 2 variable (commonly are dependent and independent variables) so we going to use chi-square and T-test which are more complex than Z score.

In unit 1 we said that you think of something a relationship between 2 variables and you have to guess the correct answer.

Example:

I am a dermatologist I wonder if there is a relationship between using sunscreen and developing skin cancer from week 1, what should I do?

First → I write down my question

Second → I write my hypothesis of correct answer (research hypothesis)

Third → I write the opposite of my hypothesis (null hypothesis)

Fourth → make some descriptive statistics

Fifth → we go from descriptive statistics that we calculate it to inferential statistics that will lead me to the answer of my research question either to keep or reject the null hypothesis

So we going to talk about the 2 main things of this inferential statistics

Inferential statistics

- Two main types:
  1) Parametric inferential statistics.
  2) Non-parametric inferential statistics.

Parametric → more preferred because the results that we get are more powerful, more flexible and more reliable than nonparametric techniques. On the other hand, to be able to use parametric test there are some assumptions and conditions that you need to check first:

1) Dependent variable (the outcome variable) should be continuous (Interval/Ratio)
2) The observations must be drawn from normally distributed populations
   If both of these assumption available then we can use it such as T-test

   Example:
   1) Relationship between using sunscreen and develop skin cancer
      The dependent variable is skin cancer but it's not continuous its yes or no variable so I can't use parametric technique
   2) Is there a relationship between gender and monthly income in dinar
      The dependent variable is monthly income and it's **continuous** (interval or ratio) and **normally distributed** so I can use a parametric test such **as T-test**

      Note: the independent variable is important in selection of the parametric test either **T-test** or **ANOVA**
      - T-test → if the independent variable has 2 categories only.
      - ANOVA → if the independent variable has 3 or more categories.

- If we cannot use parametric test and we still need to answer our research question we use non-parametric test.

  Non-parametric → the most famous one is chi-square
  - Used them when dependent and independent variables is categorical
  - Nonparametric methods require no assumptions about the population probability distributions.

- In fact both parametric and non-parametric both are valid (correct), it is only that **parametric** more powerful and flexible

Summary Table of Statistical Tests

Level of Measurement	Sample Characteristics					Correlation
	1 Sample	2 Sample		K Sample (i.e., >2)		
		Independent	Dependent	Independent	Dependent	
Categorical or Nominal	X^2	X^2	Macnarmar's X^2	X^2	Cochran's Q	
Rank or Ordinal		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	t test between groups	t test within groups	1 way ANOVA between groups	1 way ANOVA (within or repeated measure)	Pearson's r
		Factorial (2 way) ANOVA				

Memorize – Memorize—Memorize – Memorize – Memorize – Memorize – Memorize

The table is an exam question for sure (the last column is not for memorization).

Example:

1) Is there a relationship between gender and body mass index if you know that body mass index in Jordan is normally distributed?

<u>Answer:</u>

Dependent → body mass index

it is continuous and normally distributed (parametric

Independent → gender (2 sample male or female)

So we use T-test

2) If we recoded the body mass index to obesity status (in the previous example) in Jordan and the possible answer is

a) Underweight **b) Normal weight**

c) over weight **d) obese**

Answer:
Dependent → obesity status it is not continuous (it is ordinal)
Independent → Gender (2 sample)
So we use Mann Whitney U (Non-parametric test)

3) I have a sample of Jordanian some live in city and some live in village is there a relationship between poverty status and place of residence (city or village)

Answer:
Dependent → poverty status and it is yes or no question (poor or not) categorical or nominal
Independent → place of residence (2 sample)

So we use X^2 (chi-square)

4) is there a relationship between how much money you make and your specialty if your sample contain Doctors, Pharmacists and Nurse ?

Answer:
Dependent → money you make and it is continuous (parametric)
Independent → specialty (3 sample)

So we use 1 way ANOVA

Test your self

Q1/ Assessing weight loss after a nutrition intervention among the one group of 50 students who receive the intervention. Would like to determine if there is a relationship between participation in the intervention and weight loss. Weight is measured before and after the intervention?

1) T-test
2) Chi-square
3) ANOVA
4) Z test
5) Both 1 and 4 correct

Q2/ Student government on your campus sponsors a wide range of volunteer activities. You are interested in whether students who volunteer are generally more active than those who do not. You distribute a survey in randomly selected classes on campus that asks students to rate their level of involvement in campus activities using a 5-point Likert scale. You compare the ratings of volunteers to those of non-volunteers. Which of the following statistics would you use for your comparison?

1) Chi-square test.
2) Mann-whitney test
3) T-test
4) Friendman's ANOVA
5) Kruskal wallis

All thanks for Luai Shaban (2nd year medical student at turkey) because of his help in writing the sheet

Best wishes

ANSWER : 1→5 2→2