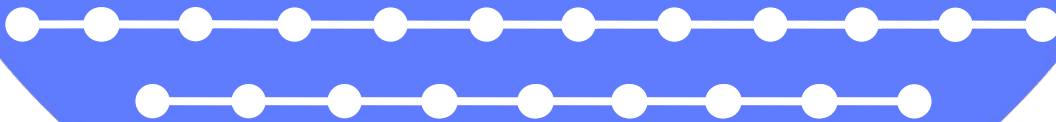




# 22

## STATISTICAL TESTS FOR QUANTITATIVE VARIABLES



KSU COLLEGE OF MEDICINE  
2019 - 2020

### ACKNOWLEDGMENTS

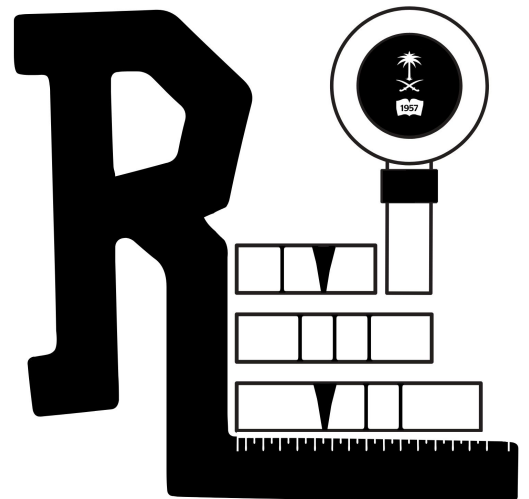
#### DONE BY

Lama Alhadlaq

Noura Alothaim

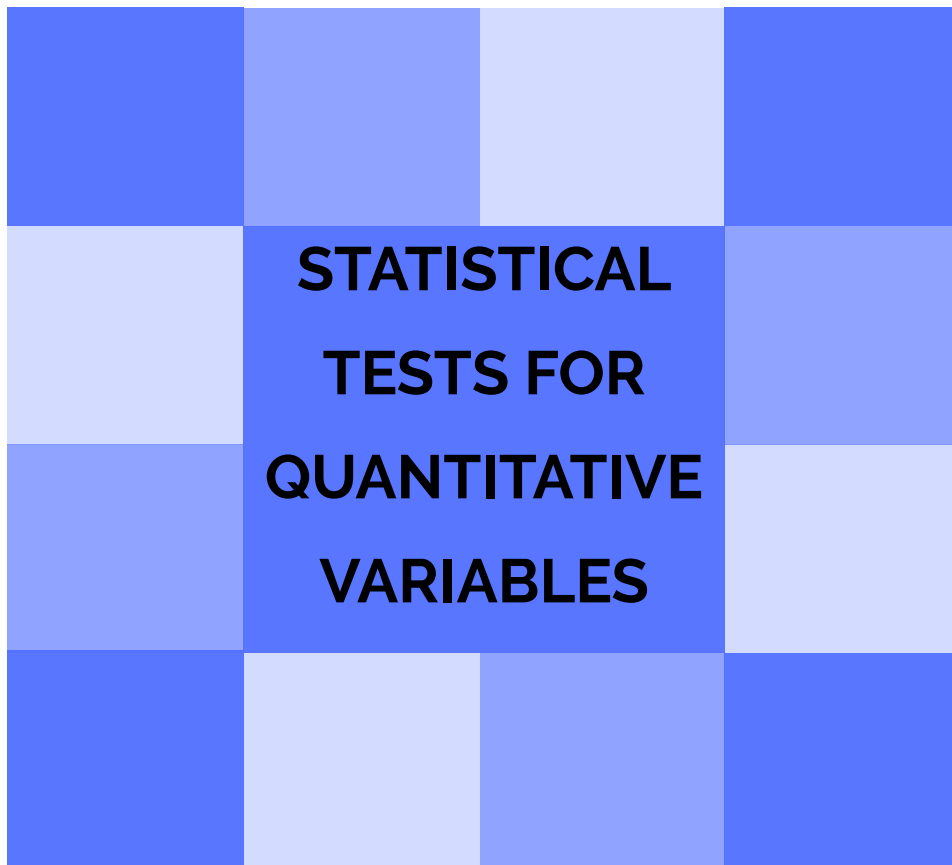
#### REVIEWER

AFNAN ALMUSTAFA



Special thanks to SARAH ALANZI & 436 TEAM

# TABLE OF CONTENTS



## LECTURE OBJECTIVES



By the end of this lecture, I am able to:

- No Objectives
- 
-

# Statistical tests

## Choosing the appropriate statistical test:

**Based on three aspects of data : IMPORTANT!!!**

1. Types of variables
2. Sample size
3. Number of groups being compared (in z test is high while in t test is low)

Test	Z-test (AKA normal distribution)	Student's t-test
Study variable	Qualitative	Qualitative
Outcome variable	Quantitative	Quantitative
Comparison	Sample mean with population mean & two sample means	sample mean with population mean; two means (independent samples); paired samples.
Sample size	larger in each group (>30) & standard deviation is known	each group <30 ( can be used even for large sample size)

- Example( Comparing Sample mean with Population mean):
  - The education department at a university has been accused of "grade inflation" in medical students with higher GPAs than students in general.
  - GPAs of all medical students should be compared with the GPAs of all students.
    - There are 1000s of medical students, far too many to interview.
    - How can this be investigated without interviewing all medical students ?

- What we know:

- The average GPA for all students is 2.70. So This value is a parameter.
- To the right is the statistical information for a random sample of medical students:

$\bar{x}$ =	3.00
s =	0.70
n =	117

# Statistical Tests

## Continue example:

- Questions to ask:
  1. Is there a difference between the parameter (2.70) and the statistic (3.00)?
  2. Could the observed difference have been caused by random chance?
  3. Is the difference real (significant)?
- Answers:
  1. The sample mean (3.00) is the same as the pop. mean (2.70). *this is null hypothesis*  
-The difference is trivial and caused by random chance.
  2. The difference is real (significant). *this is the assumption*  
- Medical students are different from all other students.

## Step 1

- Make Assumptions and Meet Test Requirements: *Hypothesis testing is a way for you to figure it out if results from a test are valid or repeatable.*
- **Random sampling**  
Hypothesis testing assumes samples were selected using random sampling.  
In this case, the sample of 117 cases was randomly selected from all medical students.
- **Level of Measurement is Ratio scale**  
GPA, the mean is an appropriate statistic.
- **Sampling Distribution is normal in shape**  
This is a "large" sample ( $n \geq 100$ ).

## Step 2

- State the Null Hypothesis  
 $H_0: \bar{x} = \mu$ 
  - You can also state  $H_0$ : No difference between the sample mean and the population parameter
  - (In other words, the sample mean of 3.0 really the same as the population mean of 2.7 – the difference is not real but is due to chance.)
  - The sample of 117 comes from a population that has a GPA of 2.7.
  - The difference between 2.7 and 3.0 is trivial and caused by random chance.
- $H_1: \bar{x} \neq \mu$ 
  - Or  $H_1$ : There is a difference between the sample mean and the population parameter
  - The sample of 117 comes a population that does not have a GPA of 2.7. In reality, it comes from a different population.
  - The difference between 2.7 and 3.0 reflects an actual difference between medical students and other students.
  - Note that we are testing whether the population the sample comes from is from a different population or is the same as the general student population.

# Statistical tests

## Step 3

- Select Sampling Distribution and Establish the Critical Region
- Sampling Distribution= Z
- Alpha ( $\alpha$ ) = .05
- $\alpha$  is the indicator of "rare" events.
- Any difference with a probability less than  $\alpha$  is rare and will cause us to reject the  $H_0$ .
- Critical Region begins at  $Z = \pm 1.96$
- This is the critical Z score associated with  $\alpha = .05$ , two-tailed test.
- If the obtained Z score falls in the Critical Region, or "the region of rejection," then we would reject the  $H_0$ .

## Step 4

- Use formula to compute the test statistic  
(Z for large samples  $\geq 100$ )

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{N}}$$

- When the Population  $\sigma$  is not known, use the following formula:

$$Z = \frac{\bar{X} - \mu}{s / \sqrt{n-1}}$$

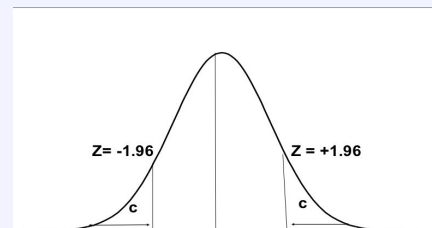
- Test the Hypotheses :  
Substituting the values into the formula, we calculate a Z score of 4.62.

$$Z = \frac{3.0 - 2.7}{.7 / \sqrt{117-1}} = 4.62$$

# Statistical tests

## Step 4

- Two-tailed Hypothesis Test:
  - When  $\alpha = .05$ , then .025 of the area is distributed on either side of the curve in area (C).
  - The .95 in the middle section represents no significant difference between the population and the sample mean.
  - The cut-off between the middle section and +/- .025 is represented by a Z-value of +/- 1.96.



## Step 5

- Make a Decision and Interpret Results
  - The obtained Z score fell in the Critical Region, so we reject the  $H_0$ .
  - If the  $H_0$  were true, a sample outcome of 3.00 would be unlikely.
  - Therefore, the  $H_0$  is false and must be rejected.
- Medical students have a GPA that is significantly different from all students ( $Z = 4.62$ ,  $p < 0.05$ ).

## Summary

- The GPA of medical students is significantly different from the GPA of all (includes non-medical) students.
- In hypothesis testing, we try to identify statistically significant differences that did not occur by random chance.
- In this example, the difference between the parameter 2.70 and the statistic 3.00 was large and unlikely ( $p < .05$ ) to have occurred by random chance.

# Statistical tests

## Example

### Weight Loss for Diet vs Exercise

Did dieters lose more fat than the exercisers?

- Diet Only:
  - sample mean = 5.9 kg
  - sample standard deviation = 4.1 kg
  - sample size = n = 42
  - standard error =  $SEM1 = 4.1/\sqrt{42} = 0.633$
- Exercise Only:
  - sample mean = 4.1 kg
  - sample standard deviation = 3.7 kg
  - sample size = n = 47
  - standard error =  $SEM2 = 3.7/\sqrt{47} = 0.540$

$$\text{measure of variability} = \sqrt{[(0.633)^2 + (0.540)^2]} = 0.83$$

- **Step 1.** Determine the null and alternative hypotheses.
  - Null hypothesis: No difference in average fat lost in population for two methods. Population mean difference is zero.
  - Alternative hypothesis: There is a difference in average fat lost in population for two methods. Population mean difference is not zero.
- **Step 2.** Sampling distribution: Normal distribution (z-test)
- **Step 3.** Assumptions of test statistic ( sample size > 30 in each group)
- **Step 4.** Collect and summarize data into a test statistic.
  - The sample mean difference = 5.9 – 4.1 = 1.8 kg (we have whether it is due to chance or because of the intervention) and the standard error of the difference is 0.83. So the test statistic:  $z = 1.8 - 0 / 0.83 = 2.17$
- **Step 5.** Determine the p-value.
  - Recall the alternative hypothesis was two-sided. p-value = 2 [proportion of bell-shaped curve above 2.17]
  - Z-test table => proportion is about 2 \* 0.015 = 0.03.
- **Step 6.** Make a decision:
  - The p-value of 0.03 is less than or equal to 0.05, so ...
  - If really no difference between dieting and exercise as fat loss methods, would see such an extreme result only 3% of the time, or 3 times out of 100.
  - Prefer to believe truth does not lie with null hypothesis. We conclude that there is a statistically significant difference between average fat loss for the two methods.

# Statistical tests

## Student's t-test types:

Compare two means together and tell you if there is any difference or not. if there is any, it will tell whether it is significant or not

1. Test for single mean :

Whether the sample mean is equal to the predefined population mean?

2. Test for difference in means:

Whether the CD4 level of patients taking treatment A is equal to CD4 level of patients taking treatment B ?

3. Test for paired observation:

Whether the treatment conferred any significant benefit ?

## 1- Test for single mean:

### 1-Questioned to be answered

Is the Mean SBP of the sample of 20 patients is 120?

N=20,  $\bar{x}$ =135, sd=5,  $\mu$ = 120

### 2-Null Hypothesis:

The mean SBP of 20 patients is 120. That is, The sample mean is equal to normal value (population mean).

### 3-Test statistics $t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$ --- t (n-1) df **this is the degree of freedom**

### 4-Comparison with theoretical value

if tab t (n-1) < cal t (n-1) reject Ho,  
if tab t (n-1) > cal t (n-1) accept Ho

### ★ Test statistics

n=20,  $\bar{x}$ =135, sd=5,  $\mu$ =120.

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}} = (135-120)/1.12 = 13.39$$

t = t .05, 19 = 2.093 Accept H0 if t < 2.093

Reject H0 if t >= 2.093

Inference : We reject Ho, and conclude that the data is providing enough evidence, that the sample mean BP is significantly higher than the normal value. **(T table [click here!](#))**

### ★ Example :

- Sample of size 25 was selected from healthy population, their mean SBP =125 mm Hg with SD of 10 mm Hg . Another sample of size 17 was selected from the population of diabetics, their mean SBP was 132 mmHg, with SD of 12 mm Hg .

-Test whether there is a significant difference in mean SBP of diabetics and healthy individual at 1% level of significance





# Karl Pearson Correlation Coefficient

## Working with two variables (parameter)

- As Age ↑ BP ↑
- As Age ↑ Cholesterol ↑
- As Height ↑
- As duration of HIV ↑
- Weight ↑
- CD4 CD8 ↓

- A number called the correlation measures both the direction and strength of the linear relationship between two related sets of quantitative variables.

- Types of correlation:

1. Positive: Variables move in the same direction

- Examples: - Height and Weight -Age and BP

2. Negative: Variables move in opposite direction

- Examples: -Duration of HIV/AIDS and CD4 CD8 -Price and Demand  
-Sales and advertisement expenditure

1-Direction: it means is it positive (direct) or negative (indirect) . We determine this by either the graph direction or the mines before the numbers.  
2-Strength: AKA magnitude which means is it weak, moderate or strong. If all variables are increasing at the same time (one perfect line!) it is strong even it is positive or negative. If there is some differences it is moderate, if it is completely different then it is weak.  
3-Linear: as one increases the other will increase and vise versa

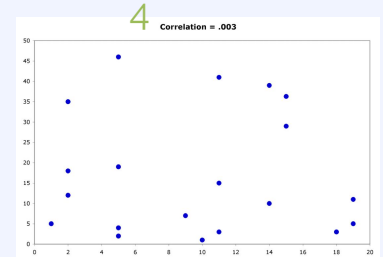
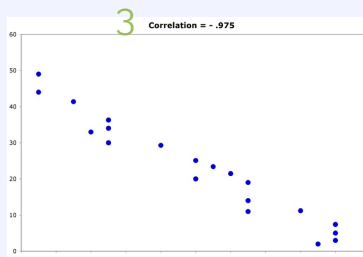
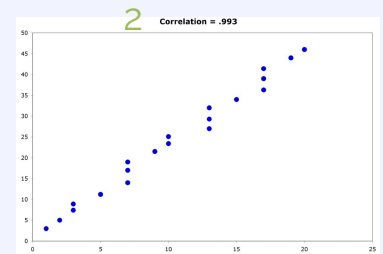
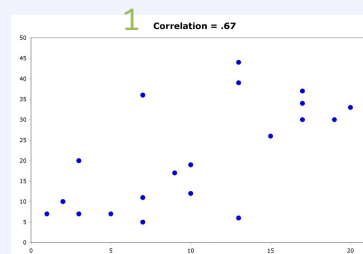
- Measurement of correlation:

- 1- Scatter Diagram
- 2- Karl Pearson's coefficient of Correlation

## Scatter diagram:

- Using the axes:
- X-axis horizontally
- Y-axis vertically
- Both axes meet: origin of graph: 0/0
- Both axes can have different units of measurement
- Numbers on graph are (x,y)

1-The direction is positive because 0.67 is positive number and it is going from down to up. The degree of strength is moderate. It is positive moderate correlation.  
2-The direction is positive because 0.993 is positive and it is going from down to up. Also, the number is closer to 1. The degree of strength is strong. It is positive perfect or strong correlation.  
3-The direction is negative because -0.975 is negative number and it is going from up to down. The degree of strength is strong. It is negative perfect or strong correlation.  
4-The direction is positive because 0.003 is positive number and it is going from down to up. the degree is weak. It is positive very weak correlation or there is no any correlation



# Karl Pearson Correlation Coefficient

The Pearson r:

- We need :
  - **Sum of the Xs**  $\Sigma X$
  - **Sum of the Ys**  $\Sigma Y$
  - **Sum of the Xs squared**  $(\Sigma X)^2$
  - **Sum of the Ys squared**  $(\Sigma Y)^2$
  - **Sum of the squared Xs**  $\Sigma X^2$
  - **Sum of the squared Ys**  $\Sigma Y^2$
  - **Sum of Xs times the Ys**  $\Sigma XY$
  - **Number of Subjects**  $(N)$

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right] \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right]}}$$

- Example:

A sample of 6 children was selected, data about their age in years and weight in kilograms was recorded as shown in the following table . Find the correlation between age and weight.

Serial No.	Age (Years)	Weight (Kg)	XY	X <sup>2</sup>	Y <sup>2</sup>
1	7	12	84	40	144
2	6	8	48	36	64
3	8	12	96	64	144
4	5	10	50	25	100
5	6	11	66	36	121
6	9	13	117	81	169
Total	41	66	461	291	742

$$r = \frac{461 - \frac{41 \times 66}{6}}{\sqrt{\left[291 - \frac{(41)^2}{6}\right] \left[742 - \frac{(66)^2}{6}\right]}}$$

r = 0.759  
strong direct correlation

- Example:

Relationship between Anxiety and Test Scores

Serial No.	Anxiety	Test score	X <sup>2</sup>	Y <sup>2</sup>	XY
1	10	2	100	4	20
2	8	3	64	9	24
3	2	9	4	81	18
4	1	7	1	49	7
5	5	6	25	36	30
6	6	5	36	25	30
Total	32	32	230	204	129

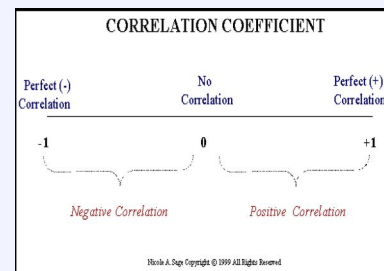
$$r = \frac{(6)(129) - (32)(32)}{\sqrt{(6(230) - 32^2)(6(204) - 32^2)}} = \frac{774 - 1024}{\sqrt{(356)(200)}} = -.94$$

r = 0.94  
Indirect strong correlation

# Correlation Coefficient

## Correlation Coefficient:

- A correlation coefficient ( $r$ ) provides a quantitative way to express the degree of linear relationship between two variables.
- Range:  $r$  is always between  $-1$  and  $1$
- Sign of correlation indicates direction:
  - high with high and low with low → positive
  - high with low and low with high → negative
  - no consistent pattern → near zero
- Magnitude (absolute value) indicates strength ( $-.9$  is just as strong as  $.9$ )
  - $0.10$  to  $0.40$  weak
  - $0.40$  to  $0.80$  moderate
  - $0.80$  to  $0.99$  high
  - $1.00$  perfect



- About "r":
  - $r$  is not dependent on the units in the problem
  - $r$  ignores the distinction between explanatory and response variables
  - $r$  is not designed to measure the strength of relationships that are not approximately straight line.
  - $r$  can be strongly influenced by outliers. *extreme values, so you must look for the extreme values and take them out.*
- Limitations:
  - Correlation coefficient is appropriate measure of relation only when relationship is linear
  - Correlation coefficient is appropriate measure of relation when equal ranges of scores in the sample and in the population.
  - Correlation doesn't imply causality:
    - Using U.S. cities a cases, there is a strong positive correlation between the number of churches and the incidence of violent crime.
    - Does this mean churches cause violent crime, or violent crime causes more churches to be built?
    - More likely, both related to population of city (3d variable -- lurking or confounding variable)
- Ice-cream sales are strongly correlated with crime rates. Therefore, ice-cream causes crime.
- Without proper interpretation, causation should not be assumed, or even implied.

**In conclusion**

- Z-test will be used for both categorical(qualitative) and quantitative outcome variables.
  - Student's t-test will be used for only quantitative outcome variables.
  - Correlation will be used to quantify the linear relationship between two quantitative variables
-