◉Sheet ○Slides

Number

9

Done by:

Sarah Al dajah

corrected by:

Mariam Hassouneh

Doctor

Mahmoud Al hussami

In today's lecture, we are going to add some new and review some old terms we took in computer skills when we were still in our 1st year :D

** We start from minute 13 , before it the doctor was talking about SPSS manual .

Proportions and Probabilities

We use probability when it's concerning the population , and we use proportions when we are concerned about a limited sample.(Proportions are usually quoted for **samples** - probabilities are almost always quoted for **populations.)**

We often interpret proportions as probabilities. If the **proportion** with a disease is 1/10 then we also say that the **probability** of getting the disease is 1/10, or 1 in 10.

- Proportions mostly are represented by percentages (EX. 2 out of 10 becomes 20%)

Workers Example

Smoking	Workers	Cases	Controls
No	Yes	11	35
	No	50	203
Yes	Yes	84	45
	No	313	270

Here we are measuring the proportions of people at risk (smokers, wither tobacco; cigarettes, chewing tobacco …etc) and controlled people who are not, to find the relationship of smoking with lung cancer.

- For the cases:
  - Proportion of exposure=84/397=0.212 or 21.2%
- For the controls:
  - Proportion of exposure=45/315=0.143 or 14.3%

Transformed into percentages!

*note: controls are <u>usually</u> not exposed to the risk of smoking (here in this example they are not smokers but they are exposed to smoke).

**The doctor didn't add details on this table

*13.00min - 17.30 min *

This leads us to other measures; prevalence and incidence .

** Remember – if you still can - what we took in community, we can also find incidence rate in the community, which is the number of new cases per population at risk in **a given time period**. If we want to find prevalence and incidence percentages, we multiply by 100.

- I found this on Wikipedia, I wish it would clarify: The difference between prevalence and incidence can be summarized as : prevalence answers "How many people have this disease right now?" and incidence answers "How many people per year newly acquire this disease?"..

- Prevalence :

Disease Prevalence = the proportion of people with a given disease at a given time.

disease prevalence =

$$\frac{\text{Number of diseased persons at a given time}}{\text{Total number of persons examined at that time}}$$

Prevalence is usually quoted as per 100,000 people so the above proportion should be multiplied by 100,000.

$$Prevalence = \frac{Cases(old + new)}{Total}$$

Old = duration of the disease
New = speed of the disease

- Incidence and prevalence rates can be found by and are used in laboratory tests (machines), and in surveys.

** 18.50min - 21.50 min ** it's better to skip this part ^^

Sensitivity and Specificity :

We are MOST concerned about the sensitivity and the specificity of the machine used in the test. In other words we are concerned about its efficiency.

- The term accuracy and efficiency are NOT used interchangeably; accuracy means the content's validity of the test and efficiency means how close the real values are to the predictive values if tested.

- Sensitivity and specificity are terms used to describe the effectiveness of screening tests. They describe how good a test is in two ways - finding false positives and finding false negatives.

Remember that these are measurements for **A SAMPLE**, and proportions are quoted for samples!

Why?

- **Sensitivity** is the **Proportion** of diseased who screen positive for the disease.
- **Specificity** is the **Proportion** of healthy who screen healthy.

	Condition Present	Condition Absent
Test Positive	True Positive (TP)	False Positive (FP)
Test Negative	False Negative (FN)	True Negative (TN)

VERY IMPORTANT NOTE : remember that the first half of these measurements' name refers to the agreement of the condition with diagnosis (FALSE means the diagnosis disagrees with the condition , TRUE means they agree) , while the second half refers to what the doctor (the diagnosis) said (positive means YES the patient is sick , negative means NO the patient is healthy).

1. Test Sensitivity (Sn) is defined as the probability that the test is positive when given to a group of patients who have the disease.
   Sn= (TP/(TP+FN))x100.

   - It can be viewed as, 1-the false negative rate.
2. The Specificity (Sp) of a screening test is defined as the probability that the test will be negative among patients who do not have the disease.
   Sp = (TN/(TN+FP))X100.

   It can be understood as 1-the false positive rate.

** The examples the doctor gave are combined in one detailed in the next page.

Positive & Negative Predictive Values:

- **The positive predictive value (PPV)** of a test is the probability that a patient who tested positive for the disease actually has the disease. PPV = (TP/(TP+FP))X 100.
  - In other words : how many times does the doctor diagnose (PREDICT) a patient as sick (say YES) and he really is ?

- **The negative predictive value (NPV)** of a test is the probability that a patient who tested negative for a disease will not have the disease. NPV = (TN/(TN+FN))X100
  - In other words : how many times does the doctor diagnose a patient as healthy (say NO) and he really is ?

These measurements are used when you are collecting your data through a machine.
WHY ? Because the results are **so related** to how efficient your machine is!
But if you were collecting it through survey then you would use other measures such as reliability and validity of the survey questions .

*note : what really **does matter** is the efficiency of the machine/test.

The Efficiency:

•The efficiency (EFF) of a test is the probability that the test result and the diagnosis agree.

•It is calculated as:

EFF = ((TP+TN)/(TP+TN+FP+FN)) X 100

- Example : The doctor gave 2 separate examples , here they are combined and enhanced for better understanding ..
We took a sample of 300 people, There were 150 person with high blood pressure (which we are aware they have a HBP) , after being tested with a machine, it turned out that only 140 are diagnosed with high blood pressure and 10 were considered healthy according to the test but **sick in real life**. For the other 150 healthy people, 145 of them the test said they were healthy while 5 were diagnosed with HBP . FIND the following : TP, TN , FP , FN , sensitivity, specificity, PPV, NPV, efficiency of the machine.

- Steps are :

1. Draw a 2×2 table , the first 4 measurements fit for each cell to fill .

		Are they diagnosed as sick?		
		YES , sick	**NO , healthy**	**Total**
Are they sick?	**YES , sick**	**140 (TP)**	**10 (FN)**	**150**
	NO , healthy	**5 (FP)**	**145 (TN)**	**150**
	Total	**145**	**155**	**300**

2. Calculate
   A. sensitivity = (140 / (140 + 10)) * 100
   B. specificity = (145 / (145 + 5)) * 100
   C. PPV = (140/ (140+ 5)) * 100
   D. NPV = (145/ (145+ 10)) * 100
   E. EFFICIENCY = ((140+ 145)/ 300)) * 100

The key is right interpreting and fitting of values inside the table!

* 17.30min - 32.00min *

- Just as mentioned in previous lectures , designs of a test are categorized into :
  1- Observational 2- experimental .

- Any phenomenon in the world should be studied in four different levels :

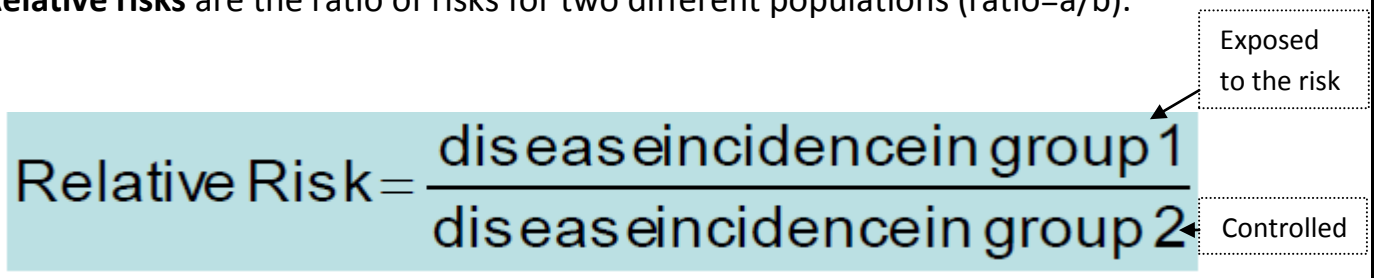
Level 1: you should explore the phenomenon and describe it (ex. For smokers, we always notice that frequent **coughing** is common between them.)
Level 2: correlation and association ; study the variables and how they are connected together
Level 3 :if the study demands and it's legally approved then you should **Manipulate** the study a bit in a way that wouldn't change the whole study and its results (done in experimental and clinical trials).
Level 4 : if manipulation wasn't advised for your study or it was **legally denied** then you resort to analytical studies and they are two types; relative risk and odd's ratio.

Before talking about relative and odd ratios you should know these terms :

- Retrospective study: A retrospective study looks backwards and examines exposures to suspected risk or protection factors in relation to an outcome that is established at the start of the study. Here lies the measuring of **Prevelance!**
- Prospective study: A prospective study watches for outcomes, such as the development of a disease, during the study period and relates this to other factors such as suspected risk or protection factor. Here lies the measuring of **INCIDENCE!**
  This study can be used after retrospective study in cases of **forbidden manipulation.**

Relative Risk and Odd's ratio:

Relative risks are the ratio of risks for two different populations (ratio=a/b).



If the risk (or proportion) of having the outcome is 2/10 in one population and 1/10 in a second population, then the relative risk is: (2/10) / (1/10) = 2.0

A relative risk >1 indicates increased risk for the group in the numerator (البسط).
A relative risk <1 indicates decreased risk for the group in the numerator.

- Relative risks are better for cohort studies (exposed and unexposed subjects are chosen and are followed to determine disease status - prospective) . It's important to know that the unexposed are not only the non- smokers for example , but also those who aren't passive smokers (ex. Wives whom their husbands smoke are not to be placed in the unexposed group or else they'll act as an extraneous variable and will give wrong interpretation).

- **Odds ratios** are better to use in case-control studies (cases and controls are selected and level of exposure is determined **retrospectively**) .

- When we have a two-way classification of exposure and disease we can approximate the relative risk by the odds ratio

		Disease		
		Yes	No	
Exposure	Yes	A	B	A+B
	No	C	D	C+D

Prospective → • Relative Risk=A/(A+B) divided by C/(C+D)

Retrospective → • Odd's Ratio= A/B divided by C/D = AD/BC

What if Odd's ratio equaled 1 ?? This means there is no relationship (A = C and B = D) , this means that whether you are exposed or not exposed , the percentage of getting the disease is the same !!!

Case Control Study Example

- Disease: Pancreatic Cancer
- Exposure: Cigarette Smoking

		Disease		
		Yes	No	
Exposure	Yes	38	81	119
	No	2	56	58

- Relative Risk= (38/119)/(2/58)=9.26
- Odd's Ratio= (38/81)/(2/56)=(38*56)/(2*81) =13.14

5 March 2017 14

Relation of cancer with smoking (Exposure):
9.26 indicates that the relationship between smokers and cancer is manifested 9 more times than that with the non-smokers .

* 32.00min – 45.56min *

Good Luck *.*