Editing file

Feedback

Appropriate statistical tests

Click here for a summary from 435

Color Index:

- Boys' Slides
- Girls' Slides
- Doctors Notes
- Golden Notes
- Important
- Extra

Extra: A quick review about the theoretical lectures from 435

Statistical tests:

3 tests for quantitative data (Z test, Student's t test and **Karl pearson correlation coefficient**)

4 test for qualitative data (Z test, Chi-square test, Mcnemar test and Fisher's exact test)

Whenever you have:

- Quantitative data (mean values)
- Qualitative/categorical (proportions)

Z test compares sample mean (or proportion) with a population mean (or proportion) or 2 sample means (or 2 proportions) (used for both quantitative and qualitative variables) sample size: >30

Student's t test is the same but only for mean values (only for quantitative) and in addition you can compare paired samples "before and after" sample size <30 and it can be also used in bigger sample size

Karl pearson correlation coefficient is where you're looking for linear relationship for two different variables (scatter diagram) Example: hours of studying and exam marks. We have positive and negative correlations.

Chi-test exclusively for qualitative (categorical) data two or more proportions (z test only 2 proportions can't compare more) sample size: >20 , expected frequency: > 5

Fisher's exact test only two proportions sample size < 20

Mcnamer's test only two dependant proportions: cross over trials & matched case control) sample size can be any size. So in the exam you have to check if they are related (dependent) or not.

Paired Samples:

For quantitative we use Student's t test

For qualitative we use Mcnamer's test

Degree of freedom (when to calculate it):

Student's t-test (one sample): n – 1

Student's t-test (two samples): n1 + n2 – 2

} **Quantitative**

Chi-square test: (R-1)x(C-1)

Fisher's Exact test: (R-1)x(C-1)

Mcnemar's test: (R-1)x(C-1)

} **Qualitative (Categorical)**

1

What are the three criteria to use, in selecting the appropriate statistical test?

Answer :
1- Type of variables.
2- Number of groups being compared.
3- Sample size.

2

One of the best indicators of the health of a baby is his/her weight at birth.
Birth weight of >2500 gms is considered normal. A researcher wants to test whether **birth weight** of babies born last year in a region are normal. He took a **sample of 100** babies and calculated **mean** and SD (Standard deviation) of the birth weights.
What test he should do to test his hypothesis that the birth weight of babies normal?

Outcome variable: Birth weight
Type of variable: Quantitative
How many groups: 1
Sample size: 100 (large) (more than 30)
Statistical test: The best test for this case is Z-test for single mean and we can also use student's t-test since its used for small and large sample size.

3

A team of scientists wants to test a new medication to see if it has either a positive or negative effect on intelligence, or no effect at all. In the population, **the average IQ** is 100 with a standard deviation of 15. **A sample of 30** participants who have taken the medication has a **mean** of 140. Did the medication affect intelligence, using alpha = 0.05? Using an appropriate statistical test they concluded that medication has significantly affected intelligence.
What is the statistical test they used here?

Outcome variable: IQ
Type of variable: Quantitative
How many groups: 1
Sample size: 30 (small)
Statistical test: Student's t-test for single mean.

Remember whenever you see mean or average its quantitative data and whenever you see proportion, out of and the frequency its categorical (qualitative) data

4

A research survey claims that 9 **out of** 10 doctors **recommend aspirin for their patients with headaches**. To test this claim, **a random sample of 100 doctors** is obtained. Of these 100 doctors, 82 indicate that they recommend aspirin. Is this claim accurate? Using an alpha of 0.05 with a two-tailed test, it was concluded that the claim that 9 out of 10 doctors recommend aspirin for their patients can't be rejected?
What is the statistical test used here?

Outcome variable: Recommend aspirin or not recommend
Type of variable: Qualitative (nominal)
How many groups: 1
Sample size: 100 (large)
Statistical test: Z-test for single proportion.

5

A statistics teacher wants to **compare** his two classes to see if they performed any differently on the tests he gave that semester. **Class A had 25 students** with an **average** score of 70, standard deviation 15. **Class B had 20 students** with an **average** score of 74, standard deviation 25. Using alpha 0.05, did these two classes perform differently on the tests?
Using an appropriate statistical test, he concluded that there was no significant difference between the performances of Class A and Class B.
What is the statistical test the teacher has used ?

Outcome variable: Score
Type of variable: Quantitative
How many groups: 2 (class A and class B)
Sample size: 20 and 25 (small) (less than 30)
Statistical test: Student's t-test for independent samples (two means).
Degrees of freedom: $n_1+n_2-2 = 20+25-2= 43$

6

Researchers want to test the effectiveness of a new anti-anxiety medication. In clinical testing, **64 out of 200 people** taking the **medication** report **symptoms of anxiety**. Of the people receiving a **placebo**, **92 out of 200** report **symptoms of anxiety**. Is the medication working any differently than the placebo? Test this claim using alpha = 0.05. what is the appropriate statistical test we can use in this situation?

Outcome variable: Symptoms of anxiety (present or absent)
Type of variable: Qualitative (nominal)
How many groups: 2 (medication and placebo)
Sample size: 200 and 200 (large) (more than 30)
Statistical test: Z-test for two proportions.

7

To test the **association** between gender and **favorite color** a study has been done on **500 college boys** and girls are asked which is their favorite color: blue, green, or pink?
Results are shown below:

	BLUE	GREEN	PINK	TOTAL
GIRLS	100	150	20	300
BOYS	20	30	180	200
TOTAL	120	180	200	500

What is the appropriate statistical test we can use in this situation?

Outcome variable: Color
Type of variable: Qualitative (nominal)
How many groups: 2
Sample size: 500 (large) (more than 20)
Statistical test: Chi-square test for independence (or association). Whenever you see association you know by default that it is Chi-Square test.
The degrees of freedom: (r-1)(c-1) = (2-1)(3-1) = 2

8

In 2010, ages of a **random sample of 500 individuals** from the same small town was taken.. below are the results:

>18 YEARS	18-35 YEARS	>35 YEARS
121	288	91

Using alpha = 0.05, would you conclude that the **population distribution of ages is equally distributed**? What is the appropriate statistical test we can use in this situation?

Is this mean or proportions ? Proportions
Outcome variable: Age
Type of variable: Qualitative (age in category) (ordinal)
How many groups: 1
Sample size: 500 (large)(more than 20)
Statistical test: Chi-square test for homogeneity (to see whether the values are distributed equally or not). In this example it isn't uniformly distributed (not homogenous because 288 which is 57.6% of 500 are in the 18-35 years category so we need to provide statistical evidence

9

Researchers want to test a new weight loss pill. The following is the **weights** (kg) of **10 people before and after** taking the pill.

BEFORE	90	100	70	50	70	50	90	60	80	70
AFTER	85	85	65	40	50	40	70	50	50	70

How to find the effect of this pill on weight loss? What test will you do in this situation using alpha = 0.05? What is the degrees of freedom?

Outcome variable: Weight
Type of variable: Quantitative
How many groups: 1
Sample size: 10 (small) (less than 30)
Statistical test: Student's t-test for paired samples (dependent samples). (because it's before and after, with small sample size)
The degrees of freedom: (n-1) = (10-1) = 9 (n-1 because its one sample)

10

When the chi-squared test for 2x2 table is not valid (when the expected numbers are <5) What is an alternative test we use?

Answer : Fisher's exact test.
Remember, the only alternative test for Chi-Square test when the table is 2x2 with a small sample size which results in an expected value less than 5 is Fisher's exact test.
Fisher's exact test is only for 2x2 table, not bigger dimensions tables.

11

A researcher wants to test the **mean systolic blood pressure** of Saudi females of Dammam city is 120 mm/hg with 95% confidence. He took a **random sample of 525 Saudi females** and found the mean systolic blood pressure as 110 mm/hg .
What is an appropriate test here to test his hypothesis?

Outcome variable: Mean systolic blood pressure
Type of variable: Quantitative
How many groups: 1
Sample size: 525 (large) (more than 30)
Statistical test: Z-test for single mean (Z test because sample size is large)

12

The following data describe numbers of children with **different sized palatine tonsils** and their carrier status for Strep. pyogenes. What is the statistical test used to **observe an association** between carrier status and size of tonsils?

	TONSILS			TOTAL
	NOT ENLARGED	ENLARGED	ENLARGED GREATLY	
CARRIERS	19	29	24	72
NOT CARRIERS	497	560	269	1326
TOTAL	516	589	293	1398

Outcome variable: Size of tonsils
Type of variable: Qualitative (ordinal)
How many groups: 2 (carriers and non-carriers)
Sample size: 1398 (large) (more than 20)
Statistical test: Chi-square test for association (or independence).
The degrees of freedom: (r-1)(c-1) = (2-1)(3-1) = 2
Can we calculate the odds ratio for this table? No, because it's 3 columns and 2 rows.

13

We wish to test the **proportion of smokers in a region is 15%**. Taking a **random sample of 320** persons in that region and **found the proportion as 22%.**
What is an appropriate test here to test the hypothesis that sample proportion is not equal to proportion of smokers in that region?

Outcome variable: Proportion of smoking (Smoker or non-smoker)
Type of variable: Qualitative (nominal)
How many groups: 1
Sample size: 320 (large) (more than 30)
Statistical test: Z-test for single proportion. We can't apply the chi-square here because its a single proportion and Chi-Square test is for 2 or more proportion.

14

A researcher wants to test the **mean HB** of a pregnant women of Malaz area is 12 g/dl. He took a **random sample of 20** and found that the mean score is 11g/dl and standard deviation is 34 g/dl. Could this sample originate from a population of mean = 12 g/dl?
What is an appropriate test here?

Outcome variable: Hemoglobin
Type of variable: Quantitative
How many groups: 1
Sample size: 20 (small) (less than 30)
Statistical test: Student's t-test for single mean. (you can't apply z-test here because sample size is small)
The degrees of freedom: (n-1) = (20-1) = 19

15

A research team claims that their new drug increase the **birth weight** of babies. In order to test this, he took a **random sample of 75** women for **treatment group** and 75 for **control group** and at the end of the study period it was found average birth Weight 3100 g and SD 420g for treatment group and for control average weight was 2750g and SD 425g.
What is an appropriate test to be done here?

Outcome variable: Birth weight
Type of variable: Quantitative
How many groups: 2 (treatment group and control group)
Sample size: 75 and 75 (large)(more than 30)
Statistical test: Z-test for two means. You can also apply student's t-test for independent samples because it can be used for small and large samples.

16

In a epidemiological survey, **1319 school children** were assessed **symptoms of severe cold** at the age of 12 and again at the age of 14. At age 12, 356 (27%) children were reported to have severe colds in the past 12 months compared to 468 (35.5%) at age 14. What test is to be used to test these proportions?

SEVERE COLDS AT AGE 12	SEVERE COLDS AT AGE 14		TOTAL
	YES	NO	
YES	212	144	356
NO	256	707	963
TOTAL	468	851	1319

Was there a significant increase of the prevalence of severe cold?

Outcome variable: Symptoms of severe cold
Type of variable: Qualitative (nominal)
How many groups: 1
Sample size: 1319 (large) (more than 20)
Statistical test: McNemar's chi-square test. (Because they are related "paired dependent sample" Follow up of the same group)(same sample at the age of 12, same sample at the age of 14)
Degrees of freedom: (r-1)(c-1) = (2-1)(2-1)= 1

17

A researcher wants to quantify the linear relationship between systolic blood pressure and age of his study subjects. What is the appropriate plot so as to observe the relationship and what statistical measure he has to apply to quantify this relationship?

Outcome variable: Systolic blood pressure and age
Type of variable: Quantitative
What is the aproperite plot: Scatter plot. By putting age on x-axis and systolic BP on y-axis.
Statistical measures: Karl pearson of correlation coefficient.

18

What is the range of correlation coefficient?

Answer : Between -1 and +1
-1 to 0 → negative correlation
0 to +1 → positive correlation
The + and - gives us the direction and the values gives us the magnitude.

19

What are the statistical tests to use, for test of association and for the measure of association?

Answer :

- **Test the association:** Chi-square test (will see either there is association or not)
- **Measure the association:** Odds ratio (for cross sectional, prospective study and case control) or relative risk (for retrospective study and RCT).

20

What are the degrees of freedom for 3 x 4 & 2 x 3 contingency tables?

Answer :
* first number is the rows, second number is the columns (Rows x Columns)

- 3 x 4 table = (3-1)(4-1) = 6
- 2 x 3 table = (2-1)(3-1) = 2

Research Team

RESEARCH TEAM
438
Keep it simple, Keep it focused

Thank you for checking our work!

Leaders:

Shuaa Khdary Sarah AlQuwayz

Abdulrhman Alsuhaibany

Member:

Haya Alanazi

Note Taker:

Leen Almadhyani Norah almasaad

Contact us:

Research4390@gmail.com