Statistics - choice of tests

Module 12 Topic 5

Current Status

- Wrong use of statistics can be seen often e.g.
  - Using a paired test for unpaired data
  - Using parametric tests when data is not normally distributed
  - Incompatibility of the test with the type of data
- Many softwares available, so performing test is easy
- Choice of test and interpretation is still an issue
- 11. Jaykaran Journal of pharmaceutical negative results.(2010)1:61

Table 16.2 Use of selected statistical methods in *Arthritis and Rheumatism* in 1967–68 and 1982, and numbers of errors found (Felson *et al.*, 1984)

	1967–68 (n = 47)	1982 (n = 74)
Statistical method:		
t test	8 (17%)	37 (50%)
Chi squared test	9 (19%)	22 (30%)
Linear regression	1 (2%)	18 (24%)
Multiple statistical tests	4 (9%)	30 (41%)
Error:		
Undefined method	14 (30%)	7 (9%)
Inadequate description of measures of location or dispersion	6 (13%)	7 (9%)
Repeated observations treated as independent	1 (2%)	4 (5%)
Two groups compared on > 10 variables at 5% level	3 (6%)	28 (38%)
Multiple *t* tests instead of analysis of variance	2 (4%)	18 (24%)
Chi squared tests used when expected frequencies too small	3 (6%)	4 (5%)
At least one of above errors	28 (60%)	49 (66%)

Selection of the test[11]

- What kind of data are we dealing with?
  - Nominal
  - Ordinal
  - Interval
  - Ratio
- Whether data is normally distributed?
- What is the aim of the study?
  - What do we want to compare?
    - Drug vs placebo?
    - Pre and post reading in the same group?
    - Incidence of a disease?
- What are the variables?

Variables

- Variables are qualities, properties, or characteristics of persons, things, or situations that change or vary and are manipulated, measured, or controlled in research.

- **More simply stated:**
  Variables are things that we measure, control, or manipulate in research.

Types of Variables

- **Independent variables** are manipulated or varied by the researcher, for example, intervention or treatment.
- **Dependent variables** are the responses, outcomes, etc. that are measured by the researcher.
- **Extraneous variables** are not part of the research design, but may have an impact on the dependent variable(s).

Levels of Measurement

- Nominal
- Ordinal
- Interval
- Ratio

Nominal-Level Variables

- Data are organized into categories
- Categories have no inherent order
- Categories are exclusive
- Categories are exhaustive
- Examples are sex, ethnicity, marital status

Examples of Nominal-Level Questions

- Do you have a loss of appetite?
- Do you smoke a lot?
- What is your ethnicity?

Ordinal-Level Variables

- Categories can be ranked in order
- Intervals between categories may not be equal
- Examples are socioeconomic status, level of education attained (elementary school, high school, college degree, graduate degree)

Examples of Ordinal-Level Questions

- Would Intervention X be your 1st, 2nd, or 3rd choice of treatment for Condition Y?
  1. First choice
  2. Second choice
  3. Third choice

- Beck Depression Scale – Sadness Item
  0 I do not feel sad
  1 I feel sad
  2 I am sad all the time and I can't snap out of it
  3 I am so sad or unhappy that I can't stand it

Interval-Level Variables

- Distances between levels of the scale are equal
- Assumed to be a continuum of values
- An example is temperature (measured in Fahrenheit or Centigrade)

Examples of Interval-Level Variables

- IQ scores
- GRE scores
- Composite scores of multi-item scales

Ratio-Level Variables

- Equal spacing between intervals
- Have an identifiable absolute zero point
- Examples are weight, length, volume, and temperature (measured in Kelvin)
- In statistical analysis, typically there is no distinction made between interval level and ratio level

Same Variable, Different Levels of Measurement

Interval level:

What is your age in years? ____

Ordinal level:

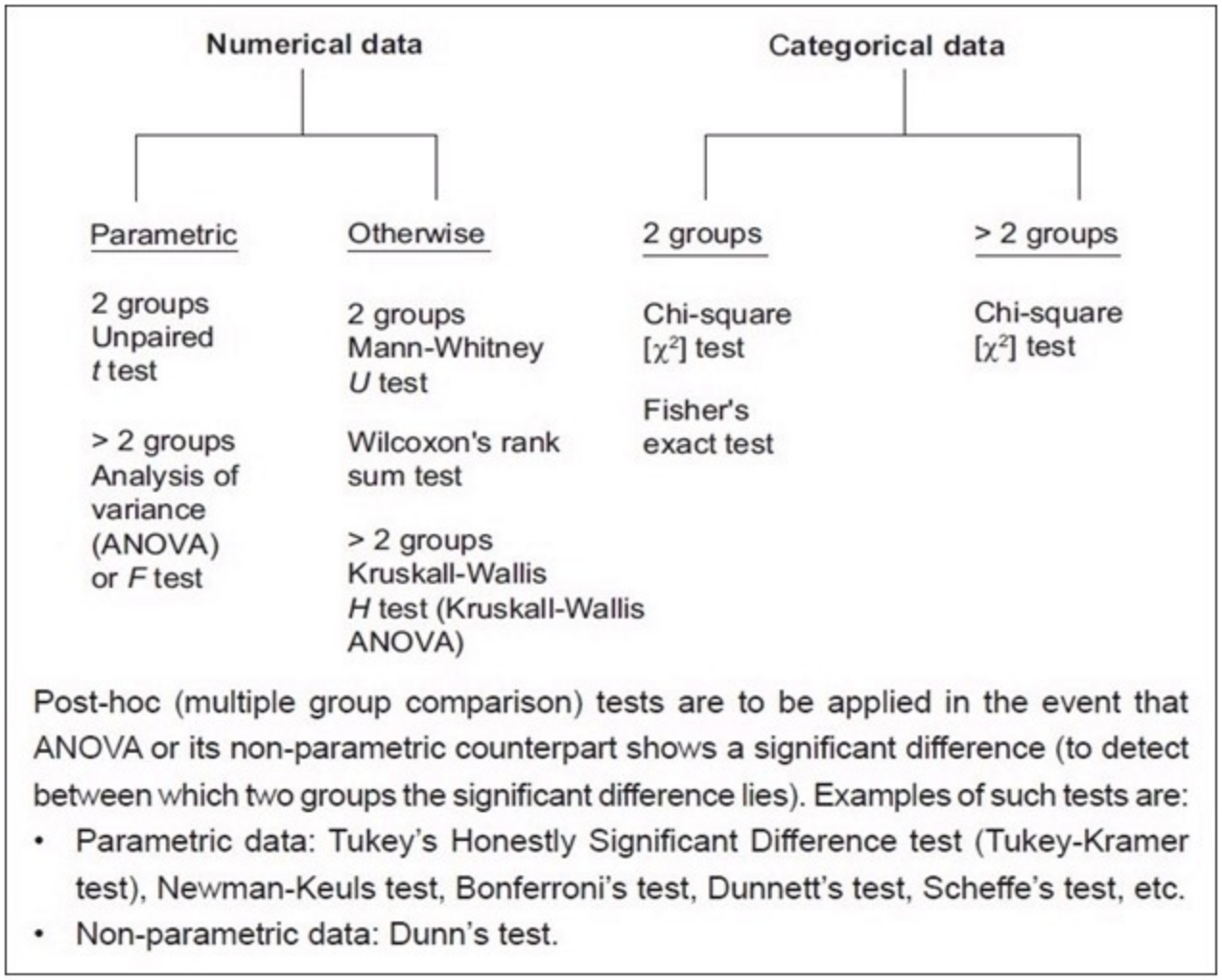
What is your age group?

- ☐ 18 years or younger
- ☐ 19-44 years
- ☐ 45 years or older

Importance of Levels of Measurement

- Level of measurement is associated with the type of statistical method used.
- Higher levels of measurement provide more information than do lower levels.
- In general, you should use the highest level of measurement possible. For example, measure actual age in years, not in age groups.

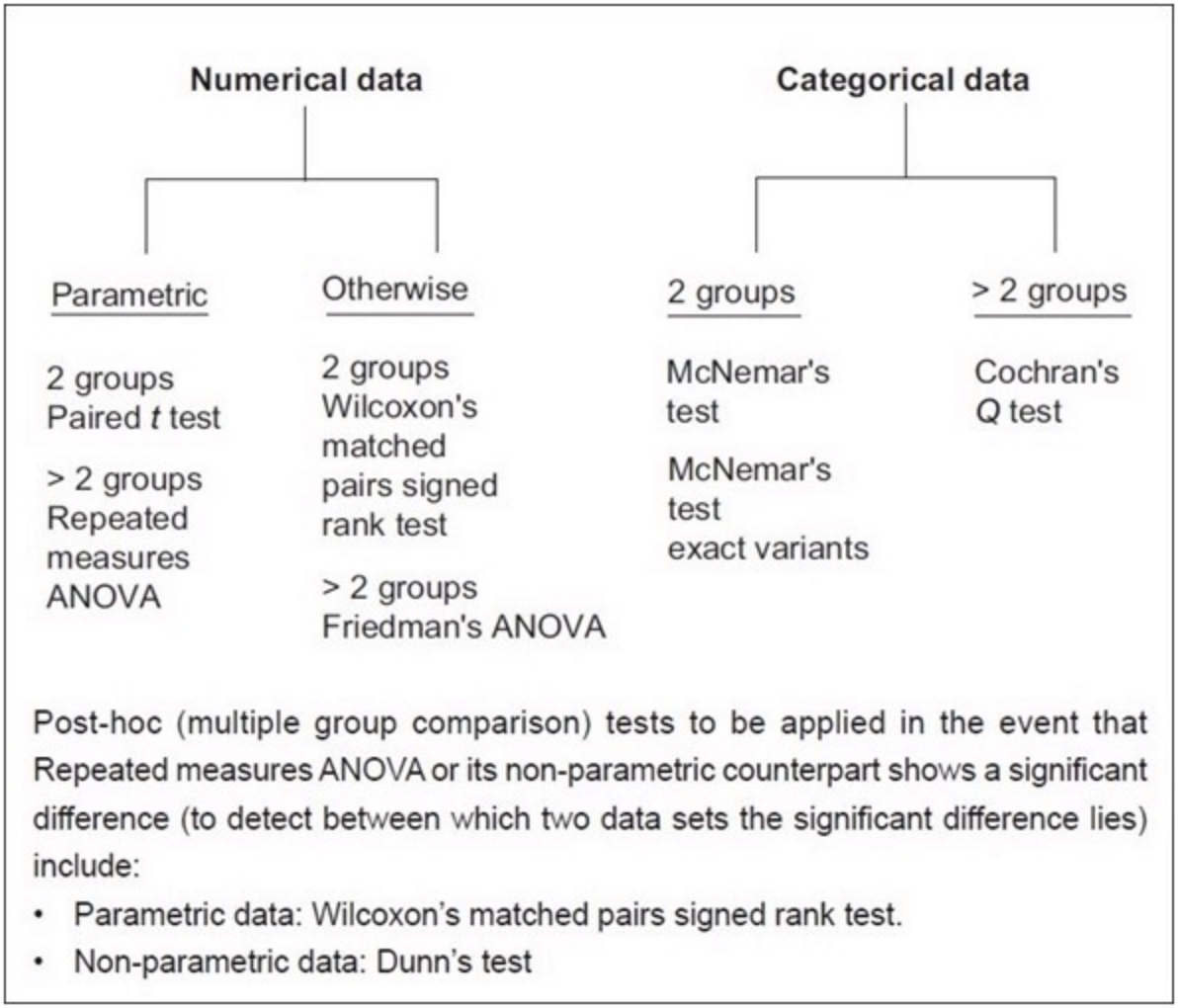
Is there a difference between groups – unpaired (parallel and independent groups) situation?



Post-hoc (multiple group comparison) tests are to be applied in the event that ANOVA or its non-parametric counterpart shows a significant difference (to detect between which two groups the significant difference lies). Examples of such tests are:

- Parametric data: Tukey's Honestly Significant Difference test (Tukey-Kramer test), Newman-Keuls test, Bonferroni's test, Dunnett's test, Scheffe's test, etc.
- Non-parametric data: Dunn's test.

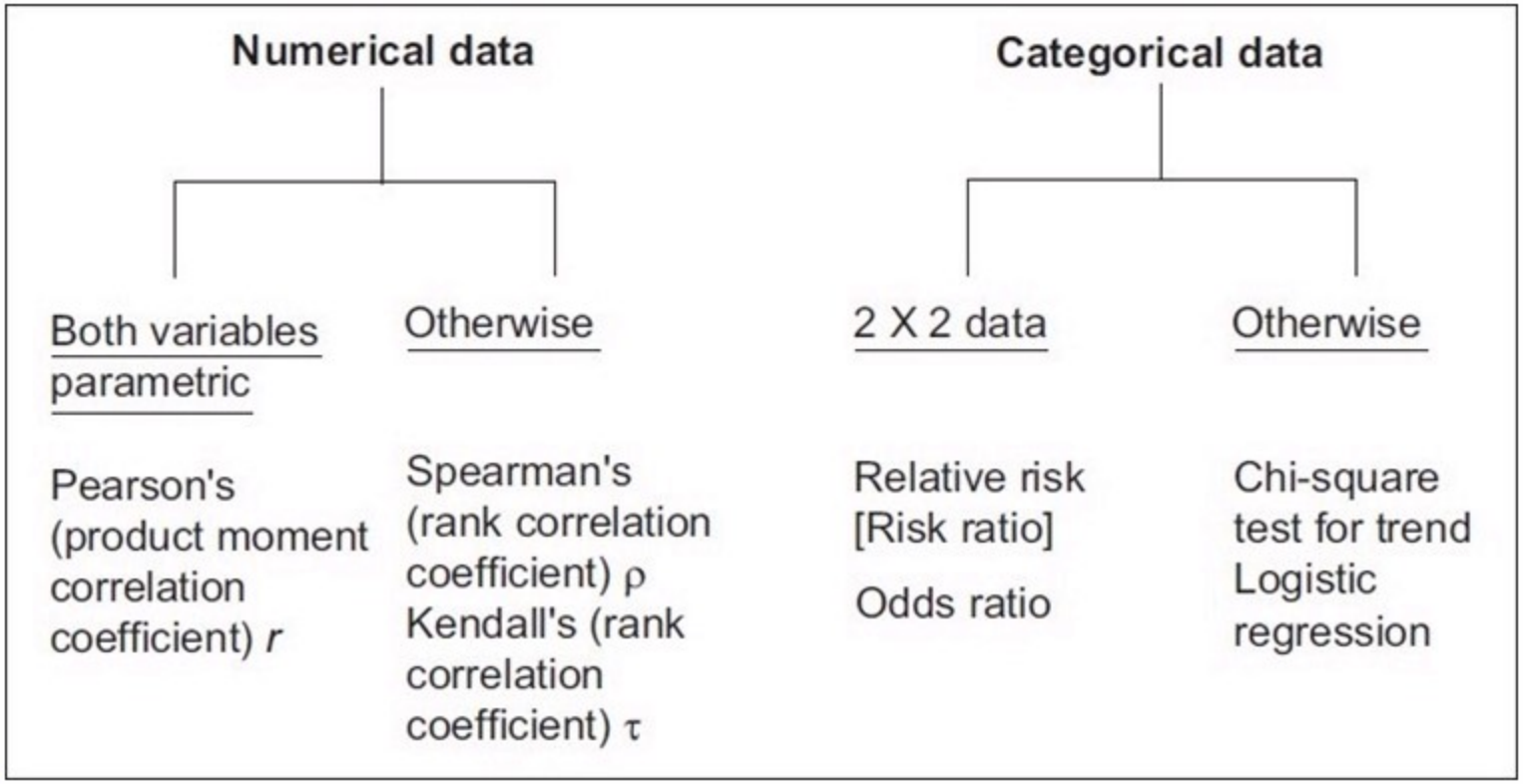
Is there a difference between groups – paired situation?(12)



Post-hoc (multiple group comparison) tests to be applied in the event that Repeated measures ANOVA or its non-parametric counterpart shows a significant difference (to detect between which two data sets the significant difference lies) include:

- Parametric data: Wilcoxon's matched pairs signed rank test.
- Non-parametric data: Dunn's test

Is there an association between variables?



Test 1

- Sex differences in student's attitudes toward homosexuals
- Hypothesis: Women are more accepting of homosexuals and thus have more positive attitudes toward homosexuals than do men
- Sample size: 10 Males and 10 females
- Measure: Scale that measures homophobia. (higher scores indicate more positive attitude)
- Responses
  - Males – 2,4,6,8,1,2,5,9,10,2
  - Women- 5,5,6,8,9,9,4,2,7,6
- Which test?
- Pearson's co-eficient, regression analysis, paired t test, unpaired t test

Test 1 - Response

- Unpaired t test - The t test for independent means (unpaired t test is used to compare means derived from unrelated /uncorrelated) samples).
- We have 2 independent groups here, women and men.
- To test a hypothesis regarding the differences between the 2 groups, we can use a t-test for independent means.
- Suppose one had continuous scores on two variables a correlation/regression could have worked. Eg machine speed vs output

Test 2

- A researcher believes that recall of verbal material differs with the level of processing. He divided his subjects into three groups.
- In the low processing group, participants read each word and were instructed to count the number of letters in the word.
- In the medium processing group, participants were asked to read each word and think of a word that rhymed.
- In the high processing group, participants were asked to read each word and try to memorize it for later recall.

Test 2

- Each group was allowed to read the list of 30 words three times, then they were asked to recall as many of the words on the list as possible.
- If the researcher wants to know whether the three groups have different amounts of recall, what type of statistical test should be used?
- Regression, one way ANOVA, unpaired t test, multiple ANOVA

Test 2 - Response

- One way ANOVA - One Way Anova can be used to test hypotheses regarding the equality of 3 or more groups. If after testing your data you found significant differences amongst your groups, you could use a post-hoc technique such as Tukey's test or even the t test to determine which specific groups performed better than others
- Two Way or Factorial ANOVA is used to test the effects of 2 independent variables. We have only 1 IV, level of processing so the Two Way test is not appropriate.

Test 2 - Response

- The t test for independent means is used to compare means derived from 2 unrelated (uncorrelated) samples. We have 3 independent groups
- Regression analysis can be used to derive an equation from which we can predict scores on one variable based on scores on another. Regression analysis is not usually used to test hypotheses about the differences between groups

Test 3

- A researcher is interested in differing interventions designed to reduce racist graffiti in an inner city area.
- He recruited 9 neighbourhoods that have the problem asked members to monitor the frequency, loca
- After four weeks of baseline recording, a six week sensitivity training workshop was begun.tion, and content of racist graffiti.

Test 3

- Community members continued to monitor the frequency, location and content of any racist graffiti during this time, yielding data for 10 consecutive weeks, 4 prior to workshops and 6 after workshops began.
- The researcher now wants to analyze his data comparing the frequency and location of the graffiti before the workshop and after the workshop. What statistical technique would you advise?

Test 3 - Response

- Repeated measures ANOVA (2way/3 way/4 way) can be used to test hypotheses regarding the equality of groups and changes in a DV over time. In this example, Repeated Measures would assess both differences in communities and changes over time with regard to the amount of graffiti present
- Chi-square is a non-parametric technique used to assess differences in categorical data (e.g. yes/no responses). Additionally, chi-square is appropriate only for research situations in which the DV is measured once. In this problem, the DV is measured 10 times.

Test 3 - Response

- Regression analysis can be used to derive an equation from which we can predict scores on one variable based on scores on another. Regression analysis is not usually used to test hypotheses about the differences between groups

Test 4

- A psychologist is interested in the relationship between job satisfaction and stress.
- Within a large corporation, the psychologist asked a random sample of workers 2 questions.
- The first question asked how satisfied workers were with their job and had them rate their satisfaction on a scale from 1 to 50.
- The second question asked how stressful they found their job in a given week. Again the workers rated their stress level on a scale from 1 to 50.
- What type of statistical test best assesses the relationship between job satisfaction and level of stress?

Test 4 - Response

- Correlation is a useful tool for describing the relationship between two paired variables, in this case job satisfaction and stress
- Analysis of CoVariance is used to remove/control for the effects of confounding variables when we have independent variables with 3 or more levels. Here we have no IV's of this type nor are we interested in removing the effects of certain variables.
- One Way Anova can be used to test hypotheses regarding the equality of 3 or more groups. One Way ANOVA is appropriate only for research situations in which we have one continuous DV and 3 or more categorical levels of a single independent variable

Test 5

- A psychologist is interested in attitudes towards the disabled. She believes that contact with someone who is disabled might have an effect on peoples' attitudes.
- To test her hypothesis, she measured attitudes toward the disables both before and after contact with an individual in a wheelchair.
- What type of statistical test should she use to determine if contact with a disabled person changes peoples' attitudes toward the disabled?

Test 5 - Response

- A paired samples t-test is used to compare data which is paired, matched or before/after. In this case we have two groups that are related, the same people were measured concerning their attitudes both before and after exposure to a disabled person. Because the groups are related, a paired samples t-test is appropriate in this situation
- While correlation is a useful tool for describing the relationship between 2 paired variables it does not serve as a test of our hypothesis
- An unpaired t test is used to compare means derived from unrelated samples. We do not have 2 independent groups

Test 6

- A psychologist is interested in conducting a study of grand juries in L.A. county to see how they compare with the demographics of potential jurors in the population.
- She wants to see if jury panels are really representative of the population.
- The first variable she examines is age.

Test 6

- The percentage of people over 65 in the population of potential jurors is 25%, but the number of people aged 65 or more who were involved in grand jury trials was 38%.
- She wants to know if the difference between people over 65 on juries is significantly different than that of the population. What test should she use?

Test 6 - Response

- The Chi-Square goodness of fit test can be used to test observed frequency data such as the percentage of individuals over 65 who serve on grand jury trials against and expected value based on the number of individuals over 65 in the population
- T test cannot be used because we do not have means, but frequency data

References

1. Aggarwal YP, Statistical Methods, concepts, application, computation. Sterlingpublishers(1986)
2. Varalakshmi V et al.Statistics – Higher Secondary, First Year. Tamilnadu Textbook Corporation 2005
3. Langley R,Practical Stats. The Chauser press 1968
4. Barkan H, Annals of Cardiac Anaesthesia (2015)18:74
5. Dunn OJ, Clark VA, Basic stats, a primer for biomed sciences. 4th Ed, 2009,John Weiley and sons inc
6. Harris M, Taylor G. Medical Statistics made easy. Ist Ed, Martin Dunitz,2003
7. Krousel wood M, clinicials guide to stats part 1 The Ochsner Journal (2006) 6:2,68
8. Kadam P ,Sample size calculation . International J of Ayurveda (2010)1:55-57
9. Fox N, Hunn A ,MathewsN, Sampling and sample size calculation . The NIHR RDS EM / YH(2009)

KLINTEC Academy

References

10. Available at http://www.biostathandbook.com/kruskalwallis.html
11. Jaykaran.Journal of pharmaceutical negative results.(2010)1:61
12. NayakBK.HazraV. Indian J Ophthalmol. (2011) 59(85-86).
13. http://wise.cgu.edu/choosemod/test61.htm