Which statistical test?

Asoka Weerakkodi[1]

Sri Lanka Journal of Obstetrics and Gynaecology 2010; **32**: 106-108

As there is a bewildering variety of statistical tests, it could be very difficult for the clinician to know which particular test one should use for a given study. Unfortunately, most books written by statisticians are not very helpful to clinicians.

We need a simple practical guide. To understand statistical tests, one needs to know about data and different data types. Which test you use depends very much on the type of data you have (rather than the type of study you have done).

Data types

For practical purposes, there are only two:

Categorical: where you count (whole numbers only). E.g. number of people, or events.

Continuous: where you measure (includes decimals). E.g. height, weight.

When it comes to different data types, Continuous is easy; there are no sub-types. 'Continuous' is continuous – that's it.

Categorical could be confusing, only because people have made it so. Statistics books describe different sub-types, which are not of much practical use (in analysis). But you may need to know these for examinations.

- *Nominal*: named groups, depending on various characteristics. One group is not 'better' than another.
  E.g. race, blood group, etc.
- *Binary*: similar to above, but only two groups. E.g. dead/alive, cured/not cured, etc.
- *Ordinal/ordered/ranked*: groups are ordered and each group has a rank. E.g. social classes, age groups, etc.
- ('qualitative', 'discrete')[1]

[Practical point: we could convert continuous data into categorical by 'grouping' them, but not vice versa][2].

[1] *Consultant Obstetrician and Gynaecologist*

E-mail: asokaweerakkody@hotmail.com

Analysis

Categorical data

When it comes to analysis, categorical data is easy. Essentially there is only one test:

Chi square test

What you need to know about the Chi square test:

1. It is a *non-parametric* test.[3]
2. You must insert *actual numbers* into the cells of the contingency table, and not percentages or proportions.
3. Number in each cell must not be less than 5; in which case you should use a variant called the *Fisher's Exact Test* (in practice, the computer packages assign this test automatically).

Continuous data

For continuous data, it is a bit more complicated; which test you use depends on certain conditions:

Data normally distributed:

For paired[4] data, use Student's *paired t-test*;

For unpaired data, use *unpaired t-test*.

For multiple groups (more than two), use *ANOVA (ANalysis Of VAriance)*

These are *parametric* tests[3].

Data not in normal distribution:

Paired data: *Wilcoxon Signed rank test*

Unpaired data: *Mann Whitney U test*

Multiple groups: *Kolmogorov-Smirnov test*

These are *non-parametric* tests.

[Practical points: If the number of observations is >100, assume normal distribution, and non-normal when it is <30.

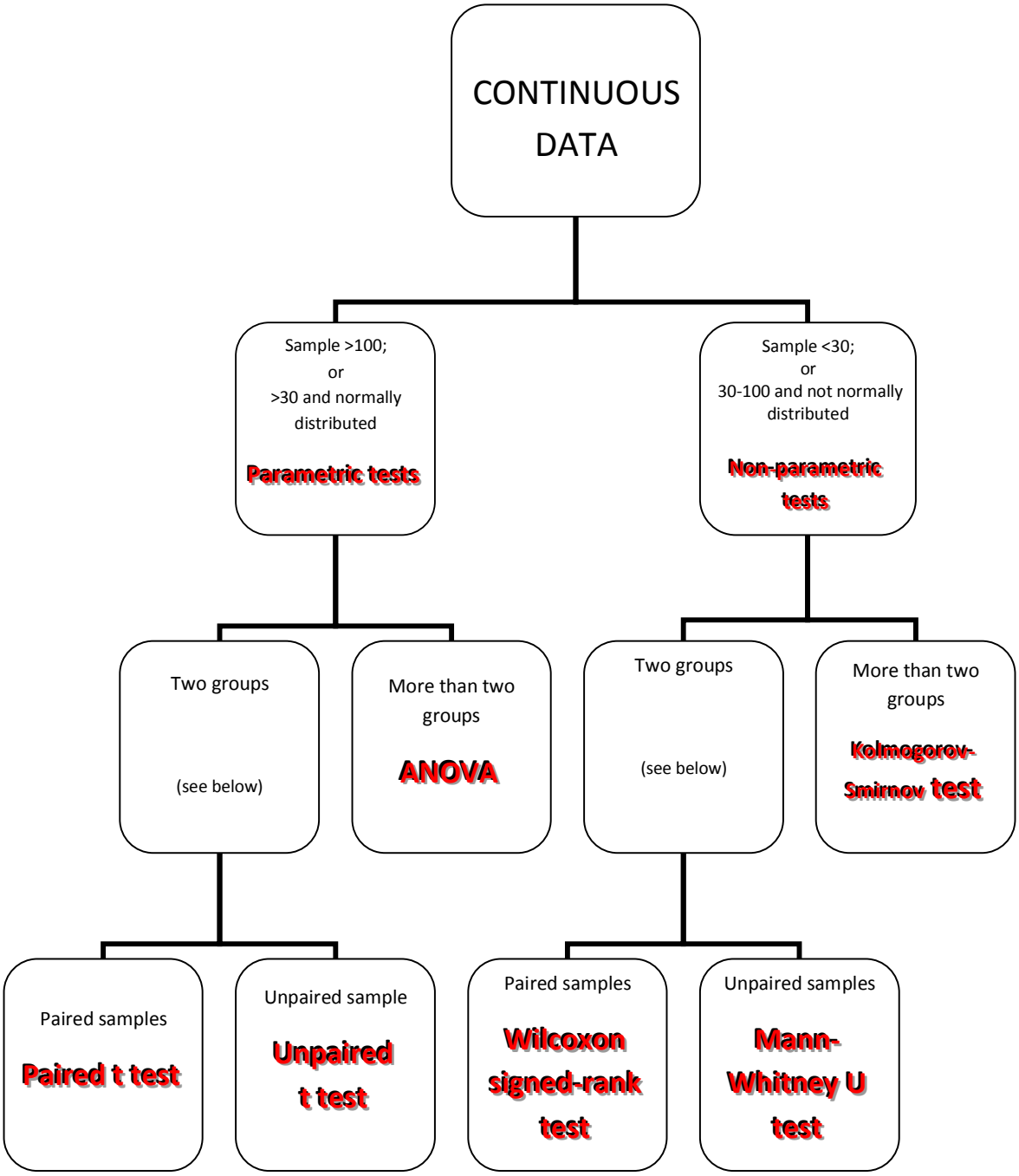
Between 30-100, one has to prove normality by various tests or by 'eye-balling'.

Sometimes it is possible to convert non-normal into normal by manipulating data, e.g. by taking the reciprocal or the logarithm.]

Summary: which test for which data?

For Categorical, use Chi Square test.

For Continuous, see chart below.



Notes

1. Some books talk about 'qualitative' data. It does not make any sense, and is best ignored. While there is a definite entity called 'qualitative research', where you describe quality, for example one's impression of the ambience of a ward, quality of care, etc. There is no such thing as 'qualitative' data in statistics. To do statistical analysis, you need quantities. (Quality can be quantified by using scoring systems and scales; but then you end up with quantities!)

   The word 'discrete' is sometimes used to describe categorical data. It is confusing and best avoided.

2. If you have a series of wt measurements, you could either enter them individually (as continuous data) or group them into 'categories' of different weight bands. In the latter case, you would analyse them as categorical data.

3. Continuous data is described by two important 'parameters' or measures, the mean and the standard deviation. The tests using these are called 'parametric'. For example, the t test could be used to compare two means and determine whether any difference observed is significant.

   All other tests, which do not use such parameters, are called 'non-parametric'.

   As a rule, parametric tests are more powerful, i.e. for the same question, they could give you the answer with a smaller sample number, for e.g. 350 vs 65!

   Thus, researchers would love to use parametric tests every time. But when applied in the wrong setting it could give quite erroneous results. That is why there are strict conditions imposed for their use, most important being the normality of the distribution.

4. Paired vs. unpaired: e.g. if you measure the vital capacity of bikers and non-bikers, you would have two sets of observations. One set is independent of the other, and they are called 'unpaired'.

   If you measure the VC of asthmatics before and after using an inhaler, the second set of readings is dependent on the first. They are called 'paired samples'.

   Paired samples are less subject to variation (as each individual acts as its own control), and hence needs less numbers than unpaired samples to show significance.

Comment

There are a large number of other statistical tests for various situations, but these are beyond the scope of the average clinician. In practice, it is unlikely that you would use, or be required to know about, anything other than the t test and the chi square test.

My sincere thanks are due to Mrs T S Usha Kiran, Consultant Obstetrician and Gynaecologist at the Royal Gwent Hospital, Newport, South Wales, UK, for valuable advice and for simplifying seemingly complex statistics.

Withdrawal from publication in the *Sri Lanka Journal of Obstetrics and Gynaecology;* Volume 32 (Supplement 1): Abstracts not presented as posters at the SLCOG annual scientific sessions in August 2010

The following abstracts published as poster presentations in the *Sri Lanka Journal of Obstetrics and Gynaecology*; Volume **32** (Supplement 1) are withdrawn. As such they should not be cited as publications.

- **PP 03: Benign intracranial hypertension (BIH) in pregnancy**
  Pathiraja PDM, Motha MBC, Wijesinghe PW
  University Obstetrics Unit, North Colombo Teaching Hospital, Ragama

- **PP 05: Heterotopic pregnancy following intrauterine insemination of donor semen followed by laparoscopic salpingectomy for the ruptured ectopic pregnancy complicated with deep vein thrombosis**
  Senaratne HMS, Dissanayake DJ, Samaranayake KU, Dissanayake TB
  Castle Street Hospital for Women, Colombo.

- **PP 06: Audit on operative vaginal delivery in a tertiary unit**
  Senaratne HMS, Dissanayake TB
  Castle Street Hospital for Women, Colombo.

- **PP 11: Disseminated endometroid adenocarcinoma of ovary in a 14-year old girl – a case report**
  Jayasundara DMCS[1], Jayalath GKC[1], Senanayake HM[1], Premarathne S[2]
  [1]*Professorial Gynaecology Unit, De Soysa Hospital for Women, Colombo 8,*
  [2]*National Cancer Institute, Maharagama.*

Editor
Sri Lanka Journal of Obstetrics and Gynaecology