

Statistics. Types of Data.

What are Statistics

The science of *assembling* and *interpreting* numerical *DATA* *Bland*

The discipline concerned with the treatment of numerical *DATA* derived from groups of individuals *Armitage*

Statistical analysis features in the majority of papers published in health care journals. Most health care practitioners will need a basic understanding of statistical principles, but not necessarily full details of statistical techniques. Medical statistics can contribute to good research by improving the design of studies as well as suggesting the optimum analysis of the results. Medical statisticians should be consulted early in the planning of a study. They can contribute in a variety of ways at all stages and not just at the final analysis of the data once the data have been collected.

Most health care practitioners do not carry out medical research. However, if they pride themselves on being up to date then they will definitely be consumers of medical research. It is incumbent on them to be able to discern good studies from bad; to be able to verify whether the conclusions of a study are valid and to understand the limitations of such studies.



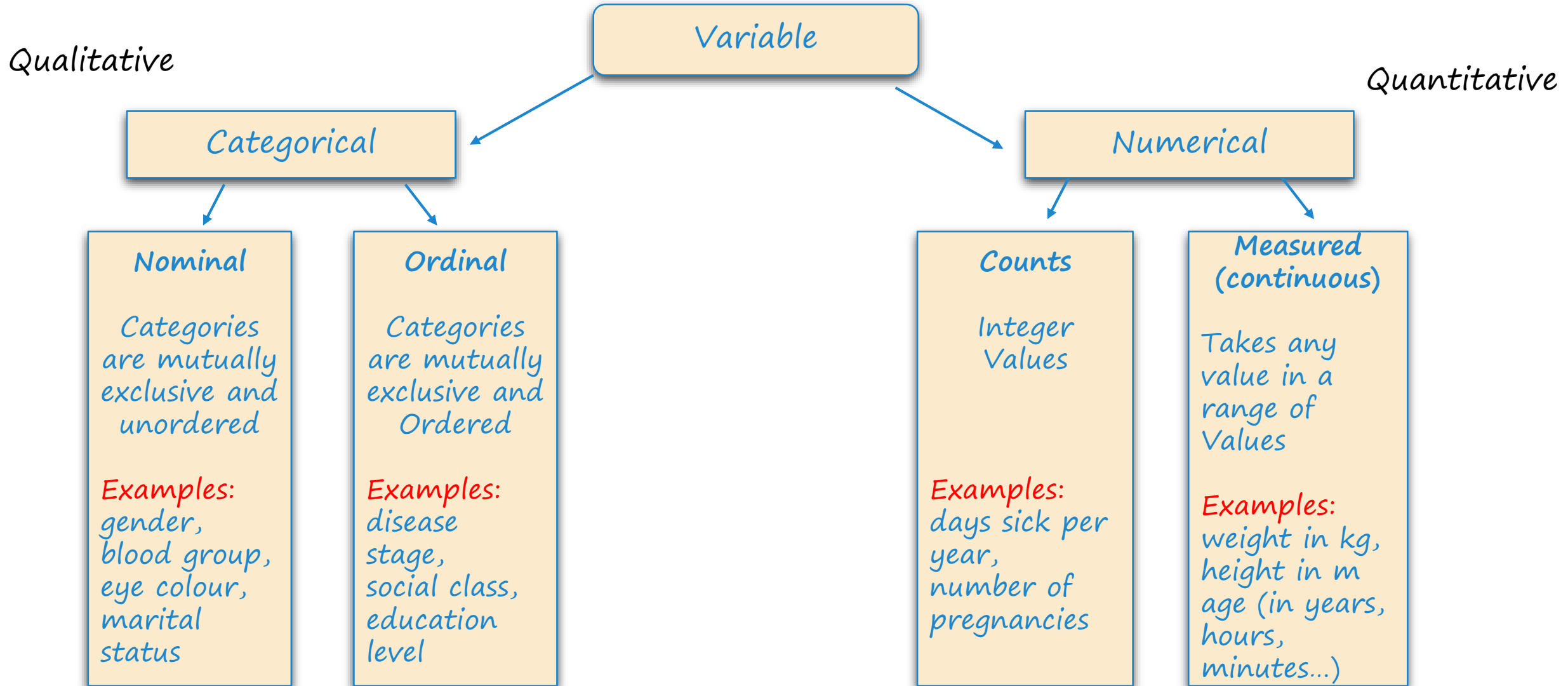
example

might be a paper describing the results of a clinical trial of a new drug.

A physician might read this report to try to decide whether to use the drug on his or her own patients. Since physicians are responsible for the care of their patients, it is their own responsibility to ensure the validity of the report, and its possible generalisation to particular patients.



Types of data



Categorical or qualitative data

- *Nominal categorical data.*

Nominal or categorical data are data that one can name and put into categories. They are not measured but simply counted. They often consist of unordered 'either-or' type observations which have two categories and are often known as binary.

- *For example:*

Dead or Alive; Male or Female

- However, categorical data often can have more than two categories, for example: *blood group O, A, B, AB*

Categorical or qualitative data

- *Ordinal data.*

If there are more than two categories of classification it may be possible to order them in some way.

- *For example, after treatment a patient may be either improved, the same or worse;*

Numerical or quantitative data

- *Count data.*

- *examples*

are often counts per unit of time such as the number of deaths in a hospital per year, or the number of attacks of asthma a person has per month.

Numerical or quantitative data

- *Measured or numerical continuous.*
- *Such data are measurements that can, in theory at least, take any value within a given range. These data contain the most information, and are the ones most commonly used in statistics.*
- *Examples of continuous data*
- *age, and body mass index.*

How a statistician can help

- *Statistical ideas relevant to good design and analysis are not easy and we would always advise an investigator to seek the advice of a statistician at an early stage of an investigation. Here are some ways the medical statistician might help.*

Sample size and power considerations

- One of the commonest questions asked of a consulting statistician is: How large should my study be?
- For a given number of patients the probability of obtaining effects of a certain size can be calculated.

Questionnaires

- the most common design was that of a crosssectional or questionnaire survey, with approximately one third of the articles classified as such.
- The responses to a questionnaire will need to be easily coded for computer analysis and a medical statistician may be able to help with this.

Choice of sample and of control subjects

- In many studies an investigator will wish to compare patients suffering from a certain disease with healthy (control) subjects. The choice of the appropriate control population is crucial to a correct interpretation of the results.

Design of study

- It has been emphasised that design deserves as much consideration as analysis, and a statistician can provide advice on design. In a clinical trial, for example, what is known as a double-blind randomised design is nearly always preferable, but not always achievable.

Laboratory experiments

- Medical investigators often appreciate the effect that biological variation has in patients, but overlook or underestimate its presence in the laboratory.

Displaying data

- A well-chosen figure or graph can summarise the results of a study very concisely. A statistician can help by advising on the best methods of displaying data.

A well-designed study, poorly analysed, can be rescued by a reanalysis but a poorly designed study is beyond the redemption of even sophisticated statistical manipulation.