CAP747:PROBABILITY AND STATISTICS

Introduction to statistics :

- ✓ different types of data
- ✓ Tables
- ✓ Charts
- ✓ histograms
- ✓ frequency distributions
- ✓ measures of central tendency: mean, median, mode, box whisker plot
- ✓ measures of dispersion: range, inter-quartile range, deviation, variance, standard deviation, standard error

Different types of data

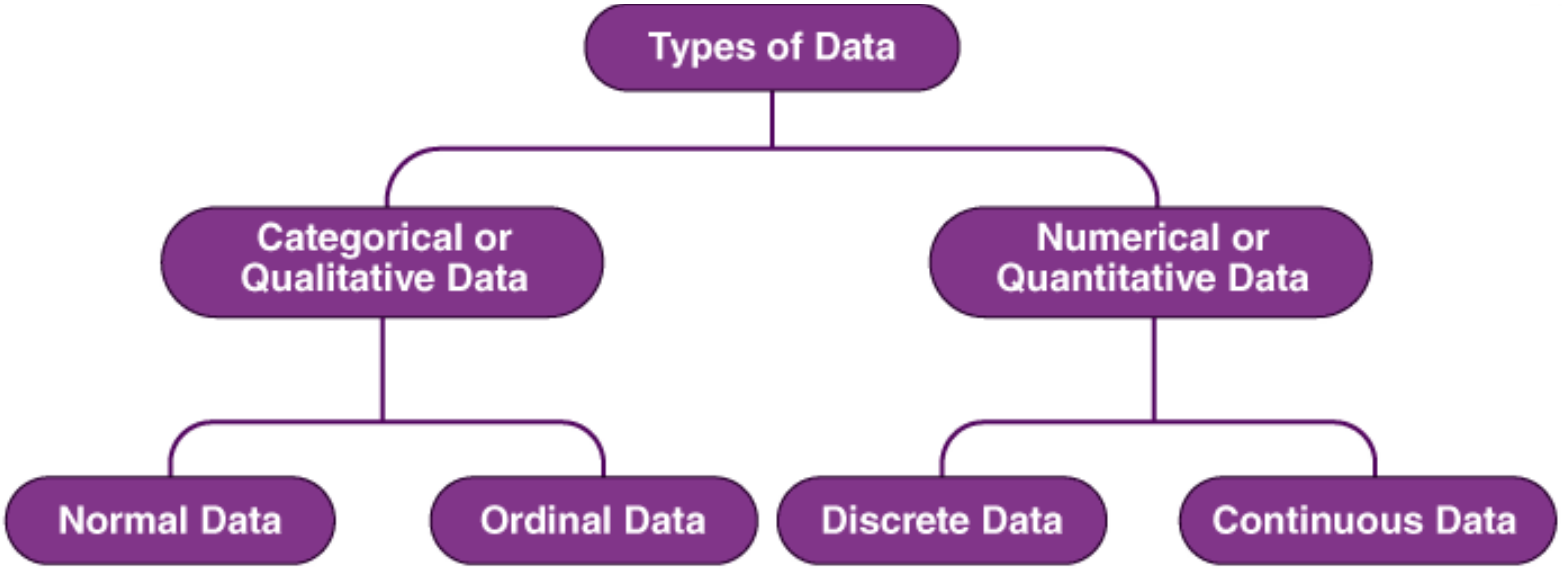
There are different types of data in Statistics, that are collected, analysed, interpreted and presented. The data are the individual pieces of factual information recorded, and it is used for the purpose of the analysis process

What are Types of Data in Statistics?

The data is classified into majorly four categories:

- Nominal data
- Ordinal data
- Discrete data
- Continuous data

Further, we can classify these data as follows:



- **1.Qualitative data**, also known as the categorical data, describes the data that fits into the categories. Qualitative data is a bunch of information that cannot be measured in the form of numbers. It is also known as categorical data. It normally comprises words, narratives, and we labelled them with names.

- It delivers information about the qualities of things in data. The outcome of qualitative data analysis can come in the type of featuring key words, extracting data, and ideas elaboration.

For examples:

- Hair colour- black, brown, red
- Opinion- agree, disagree, neutral

Question :

Is there any chance that categorical data can hold numerical values ??

Question :

Is there any chance that categorical data can hold numerical values ??

-Even if categorical variables are not quantifiable, they can appear as numbers in a data set. Correspondence between these numbers and the categories is established during data coding.

-Counts can be either of the three options: Scale (interval or ratio), categorical (ordinal, most likely), count. The third option, count, means it is a special quantitative feature called "frequency" which is similar but not identical to ratio scale. The choice is the investigator's and in a particular project.

Subdivision *of qualitative data*

1. Nominal data is one of the types of qualitative information which helps to label the variables without providing the numerical value. Nominal data is also called the nominal scale. It cannot be ordered and measured. Examples of nominal data are letters, symbols, words, gender etc.

The nominal data are examined using the grouping method. In this method, the data are grouped into categories, and then the frequency or the percentage of the data can be calculated. These data are visually represented using the pie charts.

Question:

Can nominal data be quantitative ??

2. Ordinal Data

Ordinal data is almost the same as nominal data but not in the case of order as their categories can be ordered like 1st, 2nd, etc.

Ordinal Data is observed but not measured

for example, greater or less than, higher or lower etc.

You can't do any numerical activities with ordinal data, however, as they are numerical data.

- **Quantitative data** is a bunch of information gathered from a group of individuals and includes statistical data analysis. Numerical data is another name for quantitative data. Simply, it gives information about quantities of items in the data and the items that can be estimated. And, we can formulate them in terms of numbers.

For examples:

- We can measure the height (1.70 meters), distance (1.35 miles) with the help of a ruler or tape.
- We can measure water (1.5 litres) with a jug.

Discrete Data

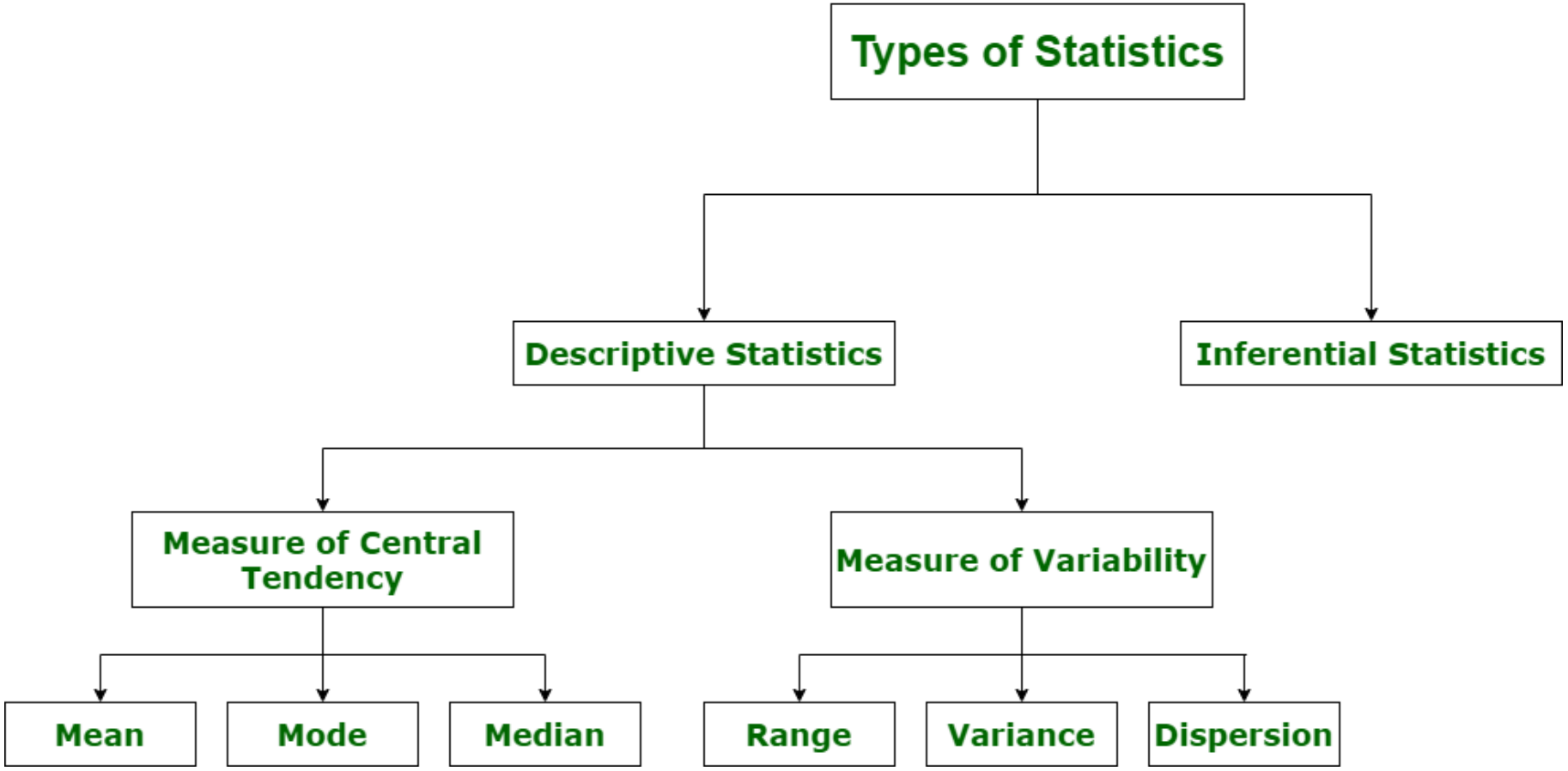
- Discrete data can take only discrete values. Discrete information contains only a finite number of possible values. Those values cannot be subdivided meaningfully. Here, things can be counted in whole numbers.

Example: Number of students in the class

Continuous Data

- Continuous data is data that can be calculated. It has an infinite number of probable values that can be selected within a given specific range.
- **Example:** Temperature range

Types of Statistics



Descriptive Statistics

Descriptive statistics uses data that provides a description of the population either through numerical calculation or graph or table. It provides a graphical summary of data. It is simply used for summarizing objects, etc.

There are two categories in this as following below.

1.Measure of central tendency

I. **Mean**

II. **Median**

III. **Mode**

2.Measure of Variability

i. Range
ii. Variance
iii. Dispersion

Inferential statistics

- **Inferential statistics** – The methods used to determine something about a population on the basis of a sample
  - **Population** –The entire set of individuals or ob
  - jects of interest or the measurements obtained from all individuals or objects of interest
  - **Sample** – A portion, or part, of the population of interest

Inferential statistics can be done with help of various steps as given below:

- Obtain and start with a theory.
- Generate a research hypothesis.
- Operationalize or use variables
- Identify or find out population to which we can apply study material.
- Generate or form a null hypothesis for these population.
- Collect and gather a sample of children from population and simply run study.
- Then, perform all tests of statistical to clarify if obtained characteristics of sample are sufficiently different from what would be expected under null hypothesis so that we can be able to find and reject null hypothesis.

Types of inferential statistics

Various types of inferential statistics are used widely nowadays and are very easy to interpret. These are given below:

- One sample test of difference/One sample hypothesis test
- Confidence Interval
- Contingency Tables and Chi-Square Statistic
- T-test or Anova
- Pearson Correlation
- Bi-variate Regression
- Multi-variate Regression

Tables(descriptive)

A **table** is an arrangement of information or data, typically in rows and columns

Types of Tables

1.General table

2. Summary tables.

1.General table

General tables contain a collection of detailed information including all that is relevant to the subject or theme. The main purpose of such tables is to present all the information available on a certain problem at one place for easy reference and they are usually placed in the appendices of reports.

Introduction to Veterinary and Comparative Forensic Medicine
John E. Cooper, Margaret E. Cooper

PART D

Appendices, References and Further Reading

2.Summary table

Summary tables are designed to serve some specific purposes. They are smaller in size than general tables, emphasize on some aspect of data and are generally incorporated within the text. The summary tables are also called **derivative** tables because they are derived from the general tables. The information contained in the summary table aims at analysis and inference. Hence, they are also known as **interpretative** tables.

Classification on the basis of characteristic

1.simple table

A simple table summarizes information on a single characteristic and is also called a univariate table.

Example

Marks of Students

Marks	0 - 10	10 - 20	20 - 30	30 – 40	40 - 50	50 - 60
Number of Students	10	12	17	20	15	6

2.complex table

A complex table summarizes the complicated information and presents them into two or more interrelated categories. For example, if there are two coordinate factors, the table is called a two-way table or bi-variate table; if the number of coordinate groups is three, it is a case of three-way tabulation, and if it is based on **more than three coordinate** groups, the table is known as higher order tabulation or a manifold tabulation.

Marks of Students

Marks	Number of Students		Total
	Males	Females	
30 – 40	8	6	14
40 – 50	16	10	26
50 – 60	14	16	30
60 - 70	12	8	20
70 – 80	6	4	10
Total	56	44	100