

STA10003

FOUNDATION OF STATISTICS

Convener/Lecturer
Mohamad Adam Bujang
019-3454379
adam@swinburne.edu.my

STA10003

Foundations of Statistics

- About this unit
 - Expectations
 - Assessment information
- Overview
 - Introduction
 - Critical thinking in statistics
 - Terminology
 - Bias
 - Summary of Module 1

STA10003

Foundations of Statistics

Learning Outcomes

After successfully completing this unit, you should be able to:

- Specify the objectives of an investigation in statistical terms.
- Identify the level of measurement of a variable, appropriate sampling methods, and research design of a study.
- Produce appropriate descriptions and visualisations of data distributions.
- Estimate population parameters (proportions, means) using appropriate statistics and confidence intervals.
- Select and conduct appropriate hypothesis tests for different population parameters (proportions, single sample means, difference of means for match and independent samples).
- Describe the relationships between variables (correlations, crosstabs, relative risk and odd ratios) and test the significance of these relationships.
- Interpret the outcomes of data analysis to write a concise report on findings.

STA10003

Foundations of Statistics

Assessments:

- Topic Tests 20% (2 tests and each carries 10%)
 - Week 5 and week 9
- Assignment 40% (2 assignments and each carries 20%)
 - Due at week 5 and week 10
- Final exam 40%
 - During examination period

STA10003

Foundations of Statistics

Tests

- Test 1 will covers from lecture at week 1 until lecture at week 4 (10%)
- Test 2 will covers from lecture at week 5 until lecture at week 7 (10%)

STA10003

Foundations of Statistics

Assignments

- Essay writing (20%)
- Mini project – Statistical analysis & interpretation (20%)

STA10003

Foundations of Statistics

Final Exam

- Final exam paper will covers lecture from all topics (40%)

STA10003

Foundations of Statistics

Note:

- Refer to the Weekly Schedule / Timeline so that you are aware of assessment due dates
- It is extremely important that you keep up with the materials in this unit –it can be difficult to ‘catch up’
- The statistical software we will use in this unit is R. It is an open-source license so you can download and install it to your own computers from <https://www.r-project.org/> and R-studio can be downloaded from <https://www.rstudio.com/> (choose the free version).

Text:

- Francis, G., & Garing, A. (2015). *Foundations of statistics*(4th ed). Frenchs Forest: Pearson Australia.

Content

- Introduction to statistics
 - a. Definition
 - b. Statistics in every day life
 - c. Critical thinking in statistics
- Terminologies in statistics
 - a. Population and sample
 - b. Descriptive and inferential
 - c. Parameter and statistics
 - d. Variables and data
 - e. Types of measurement scales
 - f. Proportion and percentage (simple statistics)

Definition of statistics

Statistics is the practice or science of collecting, managing, analyzing and interpreting data.

“Collecting”

i.e. To apply appropriate sampling techniques to get a representative sample of the targeted population

“Managing”

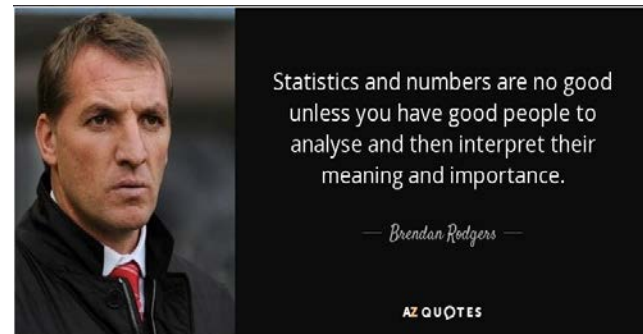
i.e. To record, clean and format the data in order to get a valid data which is feasible to be analyzed.

“Analyzing”

i.e. To analyze the data with appropriate technique to answer the objective of study.

“Interpreting data”

To convert, the statistical interpretation to common understanding.



Statistics in every day life

- Statistics is a technique to collect, manage, analyze and interpret data into significant meaning
- Statistics should not stand alone, the skills have to link and tie with other fields (social, medicine, engineering, academic, etc)
- Examples:
 - 99.9% have passed the exam
 - 57 accidents were reported this month due to traffic
 - The average age in this classroom is 19 years old
 - The intervention is effective to control glucose level (in diabetes mellitus patients)

Statistics in every day life

“Statistics is the only science that enables different experts using the same figures to withdraw different conclusions” - Evan Esar

- Think in your life that is related to statistics
- Examples?
- You will see the relation will incorporates between statistics and life
- The most important is where the conclusion should be heading to...
- This needs critical thinking!

Critical thinking in statistics

- There are many examples how statistics improve life of the people.



- Does the conclusion is correct?
- What should the better interpretation?

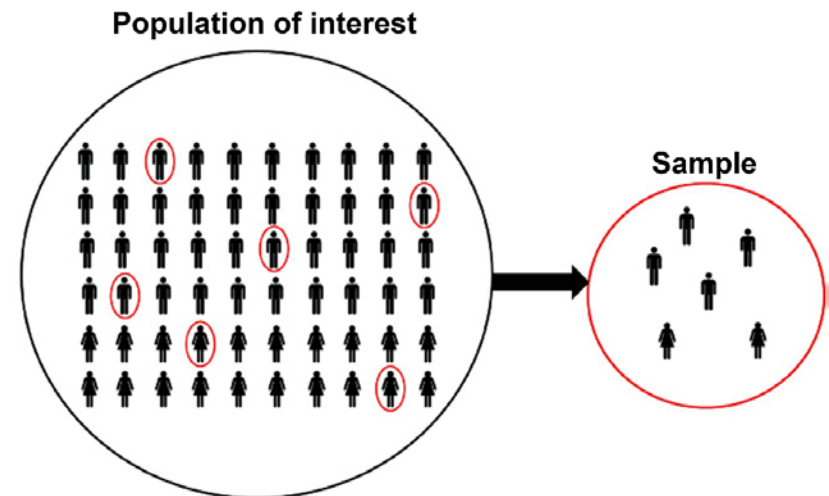
Examples:

- We're all aware that alcohol affects skill, and this finding has resulted in laws against drink driving which have led to a reduction in the road toll. The studies which established the effects of alcohol on driving skill were analysed with statistics, and it was these statistics which convinced the politicians to act.

Terminologies in statistics

Population

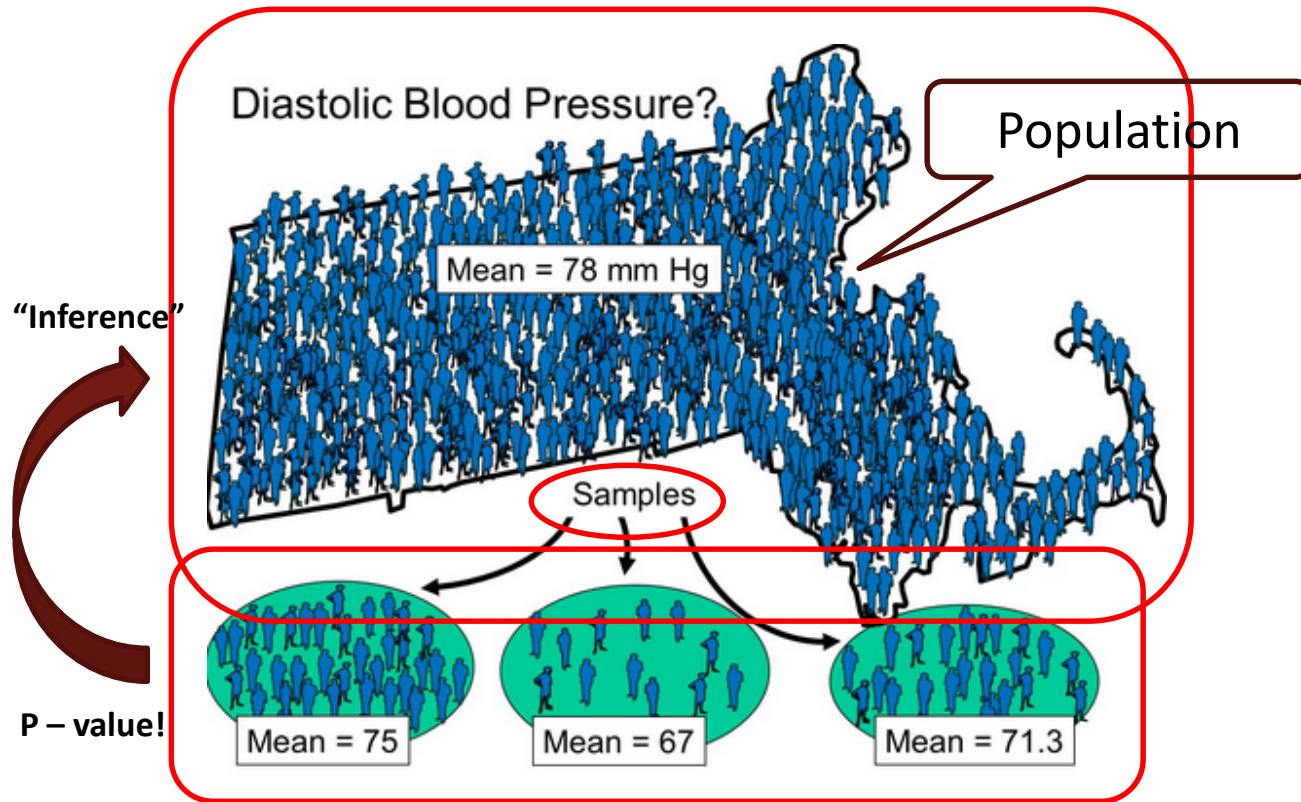
- denotes a large group consisting of elements having at least one common feature



Sample

- Subset of a population

Descriptive analysis and inferential analysis



Descriptive analysis and inferential analysis

i. Descriptive analysis

- Single data does not give much meaning.
- From a large pool of data and variables, the patterns and characteristics of the data can be observed and described.
- Describing the data actually means quantitative processes that involve describing the pattern, association and characteristics of the data. We call it as “**descriptive analysis**”.

Descriptive analysis and inferential analysis

ii. Inferential analysis

- It is usually rare researcher will collect population data to describe the pattern, association and characteristics of data.
- Due to factor of time and cost, taking sample is preferable. The question is to what extend the results from sample can represent the true value of the particular population?
- In this situation, researchers need to do inference. We call it as “inferential analysis”.
- **Inferential analysis** is a statistical technique used by researchers so they can infer the results derived from the sample to the targeted population.

Parameter and statistics

- Results derived from population data is termed as parameter
- Results derived from sample data is termed as statistics
- A statistic and a parameter are very similar. They are both descriptions of groups, like “50% of cat owners prefer Y Brand cat food.” The difference between a statistic and a parameter is that statistics describe a sample. A parameter describes an entire **population**

Variables & Data

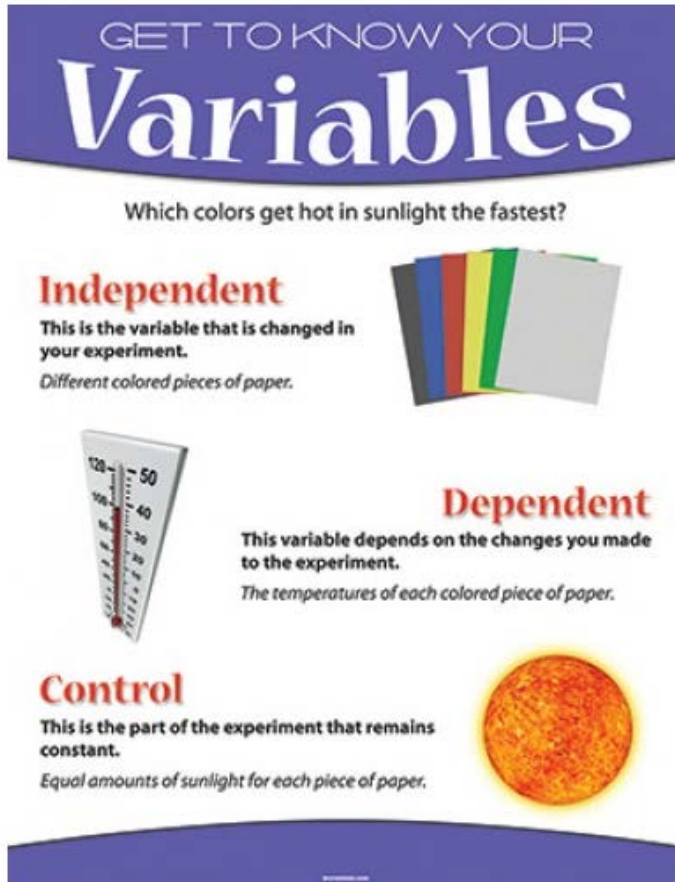
GET TO KNOW YOUR
Variables

Which colors get hot in sunlight the fastest?

Independent
This is the variable that is changed in your experiment.
Different colored pieces of paper.

Dependent
This variable depends on the changes you made to the experiment.
The temperatures of each colored piece of paper.

Control
This is the part of the experiment that remains constant.
Equal amounts of sunlight for each piece of paper.

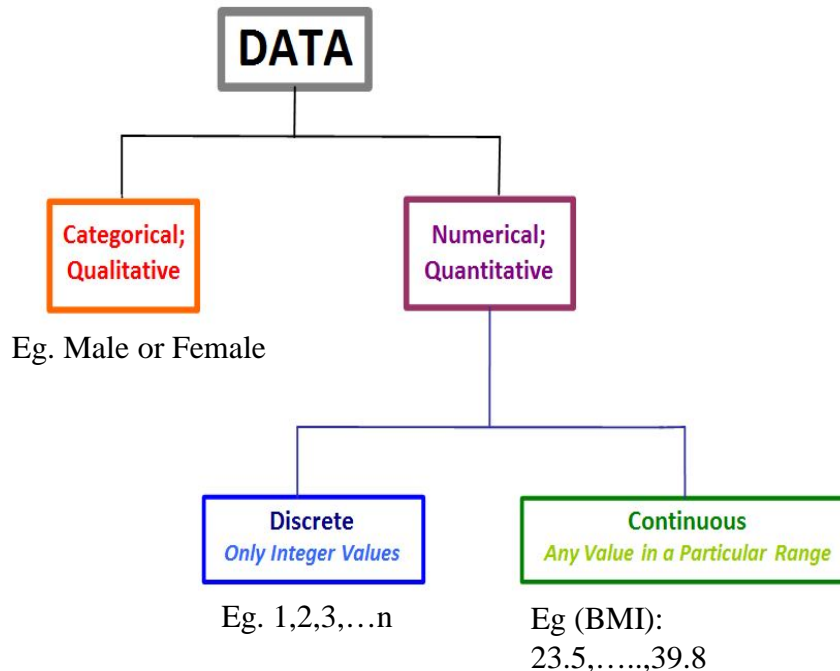


- **Variable** is something that **varies**
- Represent persons or objects or anything that can be measured
- The variables can act as independent variable (contributing factor/determinant/associated factor) or dependent variable (outcome/endpoint) or controlled variable
- Meanwhile, data is a set of values of subjects with respect to variables

Types of data

Are the data in categorical form or numerical form?

- Ethnicity
- Age
- Systolic blood pressure
- Examination mark
- Grade in exam
- Number of accidents



Proportion and percentage

- Calculating proportion: $\text{Proportion} = \frac{\text{Frequency}}{\text{Total Number}}$
 - Do first year statistics students start the course believing that statistics is just math?
 - If 90 students in a class of 140 have said that stats is maths, the proportion is
 - $90/140 = 0.643$
- Calculating Percentage
 - Proportion x 100
 - $0.643 \times 100 = 64.3\%$

Summary

- Understand why statistics is important in everyday life.
- Understand basic statistical terminologies
 - a. Population and sample
 - b. Descriptive and inferential
 - c. Parameter and statistics
 - d. Variables and data
 - e. Proportion and percentage (simple statistics)

Questions & Answers