PROBABILITY & STATISTICS

BS 1402

Grading Breakup and Policy

- Quizes: 10%
- Mid Term Exam: 30%
- Assignments/Presentations: 15%
- Class Performance: 5%
- Final Exam: 40%

Contents

- Introduction to Statistics
- The role of Probability and Statistics
- Population, Sample
- Types of Data
- Data Collection and Sampling Techniques

Introduction to Statistics

Statistics is the science of conducting studies to collect, organize, summarize, analyze, and draw conclusions from data.

- Why statistics is important?
- What are the branches of statistics?

The Role of Probability & Statistics

- Statistics is used in almost all fields of human endeavor
- Statistics is used to analyze the results of surveys and as a tool in scientific research to make decisions based on controlled experiments
- Other uses of statistics include operations research, quality control, estimation, and prediction

Population and Sample

- Population:

A **population** consists of all subjects (human or otherwise) that are being studied.

- Sample:

A **sample** is a group of subjects selected from a population.

Types of Data and Variables

- **Qualitative variables** are variables that can be placed into distinct categories, according to some characteristic or attribute.
- Examples: Gender, nationality, color etc.
- **Quantitative variables** are numerical and can be ordered or ranked.
- Examples: Age, weight, height etc.

Types of Data and Variables

Quantitative variables can be further classified into two groups: discrete and continuous.

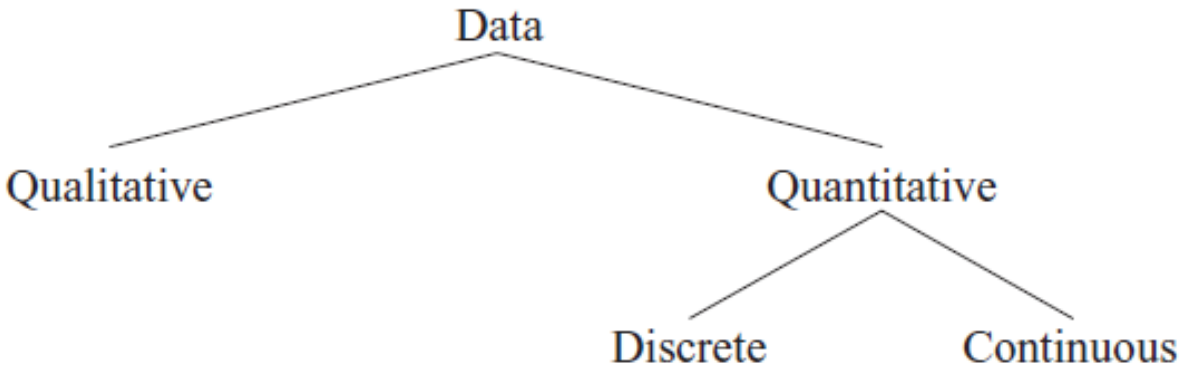
- *Discrete variables* can be assigned values such as 0, 1, 2, 3 and are said to be countable

Examples: number of students in a class, number of days in a month etc.

- *Continuous variables* can assume an infinite number of values between any two specific values. They are obtained by measuring. They often include fractions and decimals
- Examples: Height, weight, temperature etc.

Classification of Data

- Based on types of variables, data can be classified as:



Classification of Data

- Based on how data can be measured or counted, data can be classified as: nominal, ordinal, interval, and ratio

The **nominal level of measurement** classifies data into mutually exclusive (nonoverlapping), exhausting categories in which no order or ranking can be imposed on the data.

- Examples are political party (Democratic, Republican, Independent, etc.), religion (Christianity, Judaism, Islam, etc.), and marital status (single, married, divorced, widowed, separated)

Classification of Data

The **ordinal level of measurement** classifies data into categories that can be ranked; however, precise differences between the ranks do not exist.

- Examples of ordinal data are letter grades (A, B, C, D, F).

The **interval level of measurement** ranks data, and precise differences between units of measure do exist; however, there is no meaningful zero.

- This level differs from the ordinal level in that precise differences do exist between units. For example, many standardized psychological tests yield values measured on an interval scale. There is a meaningful difference of 1 point between an IQ of 109 and an IQ of 110.

Classification of Data

The **ratio level of measurement** possesses all the characteristics of interval measurement, and there exists a true zero. In addition, true ratios exist when the same variable is measured on two different members of the population.

- For example, if one person can lift 200 pounds and another can lift 100 pounds, then the ratio between them is 2 to 1. Put another way, the first person can lift twice as much as the second person.

Classification of Data

Here's a summary of classification of data based on measurement scales:

Table 1–2 Examples of Measurement Scales

Nominal-level data	Ordinal-level data	Interval-level data	Ratio-level data
Zip code	Grade (A, B, C, D, F)	SAT score	Height
Gender (male, female)	Judging (first place, second place, etc.)	IQ	Weight
Eye color (blue, brown, green, hazel)	Rating scale (poor, good, excellent)	Temperature	Time
Political affiliation	Ranking of tennis players		Salary
Religious affiliation			Age
Major field (mathematics, computers, etc.)			
Nationality			

Data Collection and Sampling Techniques

- Data can be collected in a variety of ways. One of the most common methods is through the use of surveys. Surveys can be done by using a variety of methods. Three of the most common methods are the telephone survey, the mailed questionnaire, and the personal interview.
- To avoid bias, statisticians use four basic types of sampling to collect data: random, systematic, stratified, and cluster sampling

Data Collection and Sampling Techniques

- Random Sampling:

Random samples are selected by using chance methods or random numbers.

- Systematic Sampling:

Researchers obtain systematic samples by numbering each subject of the population and then selecting every *kth* subject. For example, a researcher may select every tenth item from an assembly line to test for defects

Data Collection and Sampling Techniques

- Stratified Sampling:

Researchers obtain stratified samples by dividing the population into groups (called strata) according to some characteristic that is important to the study, then sampling from each group.

- Cluster Sampling:

Researchers also use cluster samples. Here the population is divided into groups called clusters by some means such as geographic area or schools in a large school district, etc. Then the researcher randomly selects some of these clusters and uses all members of the selected clusters as the subjects of the samples

Data Collection and Sampling Techniques

- Summary of Sampling Methods:

Table 1–4	Summary of Sampling Methods
Random	Subjects are selected by random numbers.
Systematic	Subjects are selected by using every *k*th number after the first subject is randomly selected from 1 through *k*.
Stratified	Subjects are selected by dividing up the population into groups (strata), and subjects are randomly selected within groups.
Cluster	Subjects are selected by using an intact group that is representative of the population.

Thank You.