Introduction to Probability and Statistics

Lecture 1

What is Statistics?

- Statistics answers questions using data or information about a situation
- Statistic is a property of data (eg, average)
- Statistics is the art and science of extracting answers from data

Why study statistics

- Help decision making in an uncertain environment.
- We collect and analyze data to make decisions
- We want to statements based on sample that will have some validity about the population.

Population and sample

- Population is a complete set of all items that interest an investigator. Population size N can be very large (even infinity)
- A sample is an observed subset of the population of size n.
- How are samples chosen?
- What is a Random sample?

Parameter and Statistic

- Parameter is a characteristic of the population
- Statistic is a specific characteristic of a sample

Simple exercise

- An airline claims that less than 5% of its flights from Delhi airport depart late. From a sample of 100 flights 6 flights were found to depart late
- What is the population? What is the sample? What is the statistic? Is 3% a parameter or a statistic?

Descriptive and inferential statistics

- Descriptive statistics include graphical and numerical procedures that are used to summarize data and to transform data into information
- Inferential statistics provides bases for forecast, predictions and estimates and are used to transform information into knowledge

Example - Descriptive

The number of customers who visited a jewellery shop in the last 10 days were 83, 80, 79, 85, 84, 106, 111, 120, 74, 77

Not much variation in the first five days. High next two days (weekend?). High on the next day (specific occasion?)..

Making inferences

- Estimating a parameter – average age of customers
- Test a hypothesis – weekend sales are higher than weekday sales
- Forecast sale for next month

Compare

- 6’ 3” boy to a 5’ 1” boy
- CGPA of 9.4 to CGPA of 7.8
- Income of 24 lakhs to income of 8 lakhs
- Mercedes Benz car to Alto
- 70 year old woman to 20 year old woman
- Minister to a professor

Infer or interpret

- 6’ 3” boy to a 5’ 1” boy - taller
- CGPA of 9.4 to CGPA of 7.8 – more intelligent
- Income of 24 lakhs to income of 8 lakhs - rich
- Mercedes Benz car to Alto - affordability
- 70 year old woman to 20 year old woman - health
- Minister to a professor - power

Answer questions

- How do I price this car (or air ticket)?
- How much the customer is willing to pay?
- Where should my admission cut off be?
- How tough should my question paper be?
- When should I offer a discount?
- What should be the capacity of the plant?
- How much to advertise in the world cup?

Two more aspects

- Variation in data
  - Height, weight, education, affordability, health, wealth, intelligence, abilities
- Dependencies resulting in model building
  - Linear, complex, non linear
  - Requires different data

Relevant example 1

- Data on planning MBA interviews in Mumbai?
- (IITB vs catering); timing; number of days, etc

Example 2

- Which institute(s) to apply for MBA
- What data are required?

Business Example

- Conduct IPL?
- Data required?
- Understand dependencies
- Player auctions?