Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

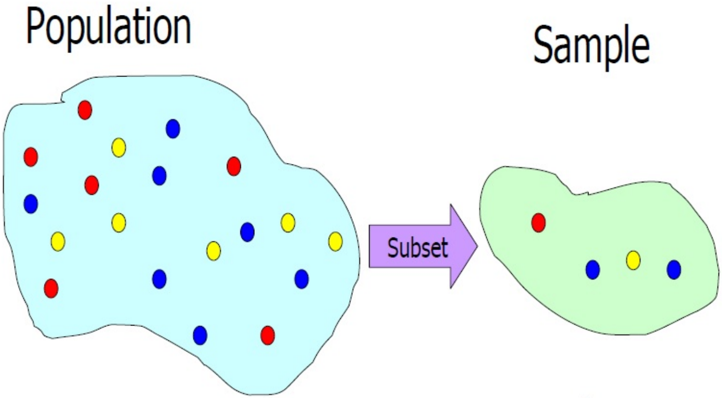
August 17, 2022

Table of Contents

Statistics is the discipline that concerns the collection, organization, displaying, analysis, interpretation and presentation of data.

Population:
In statistics, a population is a set of similar items or events which is of interest for some question or experiment.

Sample:
A sample is a subset of the population.


Population
Sample
Subset

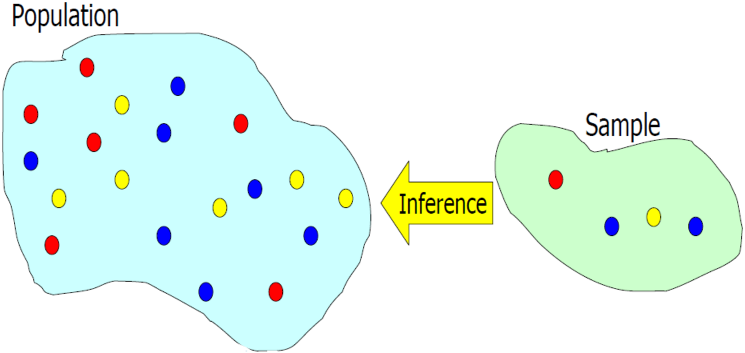
Two main branches of statistics:

Descriptive statistics: Summarize data from a sample using indexes such as the mean or standard deviation.

Two main branches of statistics:

Descriptive statistics: Summarize data from a sample using indexes such as the mean or standard deviation.

Inferential statistics:
Inferential statistics is used to make predictions or comparisons about a larger group (a population) using information gathered about a small part of that population.


Population
Inference
Sample

Sir Ronald A Fisher
(17 February 1890 – 29 July 1962)

Prasanta C. Mahalanobis
(29 June 1893– 28 June 1972)

Table of Contents

Probability Theory:

- Probability of Events, Conditional Probability and Bayes Theorem
- Random Variables and Distribution Functions
- Moments of Random Variables and Chebychev Inequality
- Some Special Discrete Distributions, Some Special Continuous Distributions
- Bivariate Random Variables and Product Moments of Bivariate Random Variables
- Some Special Discrete Bivariate Distributions, Some Special Continuous Bivariate Distributions
- Functions of Random Variables and Their Distribution, Laws of Large Numbers, The Central Limit Theorem

Statistics:

- Sampling Distributions
- Estimators of Parameters
- Test of Statistical Hypotheses
- Linear Regression

Recommended Books:

- Prasanna Sahoo, Probability and Mathematical Statistics
- Miller & Miller, John E. Freund's Mathematical Statistics with Applications
- John A. Rice, Mathematical Statistics and Data Analysis
- Larry Wasserman, All of Statistics: A Concise Course in Statistical Inference

Table of Contents

RANDOM EXPERIMENT: An experiment whose outcomes can not be predicted with certainty.

RANDOM EXPERIMENT: An experiment whose outcomes can not be predicted with certainty.

SAMPLE SPACE: A sample space of a random experiment is the collection of all possible outcomes.

RANDOM EXPERIMENT: An experiment whose outcomes can not be predicted with certainty.

SAMPLE SPACE: A sample space of a random experiment is the collection of all possible outcomes.

EXAMPLES:

1. The sample space for the possible outcomes of one toss of a coin is

$$S_1 = \{H, T\}$$

where H and T stand for head and tail.

2. If we toss it twice then

$$S_2 = \{HH, HT, TH, TT\}$$

2. If we toss it twice then

$$S_2 = \{HH, HT, TH, TT\}$$

3. If a coin is tossed until a head appears for the first time, this

could happen on the first toss, the second toss, the third toss, the fourth toss, . . ., and there are infinitely many possibilities. For this experiment we obtain the sample space

$$S_3 = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$$

with an unending sequence of elements.

Discrete Sample Space:

If a sample space contains a finite number of elements or an infinite though countable number of elements, it is said to be discrete.

Discrete Sample Space:

If a sample space contains a finite number of elements or an infinite though countable number of elements, it is said to be discrete.

- Sample spaces S_1, S_2, and S_3 (given above) are discrete.

Discrete Sample Space:

If a sample space contains a finite number of elements or an infinite though countable number of elements, it is said to be discrete.

• Sample spaces S_1, S_2, and S_3 (given above) are discrete.

Continuous Sample Space: If a sample space consists of a continuum, such as all the points of a line segment or all the points in a plane, it is said to be continuous.

• Continuous sample spaces arise in practice whenever the outcomes of experiments are measurements of physical properties, such as temperature, speed, pressure, length that are measured on continuous scales.

EXAMPLE: Choosing a point from the interval $(0, 1)$. The sample space

$$S = (0, 1)$$

is continuous.

EVENT: An event is a subset of a sample space.

EVENT: An event is a subset of a sample space.

EXAMPLE: Tossing a die. The sample space is
$S = \{1, 2, 3, 4, 5, 6\}$.$E = \{2, 4, 6\}$ is an event, which can be described in words as "the number is even".

EVENT SPACE: A subset E of the sample space S is said to be an event if it belongs to a collection $\mathcal{F}$ of subsets of S satisfying the following three rules:

(a) $S \in \mathcal{F}$

(b) If $E \in \mathcal{F}$ then $E^c \in \mathcal{F}$

(c) If $E_j \in \mathcal{F}$ then $\bigcup_{j=1}^{\infty} E_j \in \mathcal{F}$.

The collection $\mathcal{F}$ is called an event space or a σ-field.

PROBABILITY MEASURE: Let S be the sample space of a random experiment. A probability measure $P : \mathcal{F} \to [0, 1]$ is a set function which assigns real numbers to the various events of S satisfying

(a) $P(E) \geq 0$ for all event $E \in \mathcal{F}$

(b) $P(S) = 1$

(c) If $E_1,\ E_2, \ldots, E_k, \ldots$ are mutually disjoint events of S, then

$$P(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} P(E_j).$$

PROPERTIES OF PROBABILITY MEASURE:

- $P(\phi) = 0$
- $P(E^c) = 1 - P(E)$
- $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

PROPERTIES OF PROBABILITY MEASURE:

- $P(\phi) = 0$
- $P(E^c) = 1 - P(E)$
- $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

If A is an event in a discrete sample space S, then $P(A)$ equals the sum of the probabilities of the individual outcomes comprising A.

EXAMPLE: A die is loaded in such a way that each odd number is twice as likely to occur as each even number. Find $P(G)$, where G is the event that a number greater than 3 occurs on a single roll of the die.

If an experiment can result in any one of n different equally likely outcomes, and if of these m outcomes together constitute event A, then the probability of event A is

$$P(A) = \frac{m}{n}.$$

If an experiment can result in any one of n different equally likely outcomes, and if of these m outcomes together constitute event A, then the probability of event A is

$$P(A) = \frac{m}{n}.$$

EXAMPLE: If we twice flip a balanced coin, what is the probability of getting at least one head?