UNIT-I

# INTRODUCTION TO SAMPLING AND PROBABILITY THEORY

## Sampling:

Probability sampling is based on the fact that every member of a population has a known and equal chance of being selected.

~~Probability: Is the~~

**Measure of central tendency:** It is a single value that attempts to describe a set of data by identifying the central position within that set of data.

Sometimes it is also called measure of central location.

They are also classed as summary statistics.

The mean, median & mode are all valid measure of central tendency.

### 1) Mean:

It is equal to the sum of all the values in the data set divided by the total number of values in the data set when we are dealing with discrete random variable.

If we have n values in a dataset and they have values $x_1, x_2 \ldots x_n$ the sample mean, usually denoted by $\bar{x}$ is

$$\bar{x} = \frac{x_1+x_2+x_3+\ldots+x_n}{n} = \sum_{k=1}^{n} \frac{x_k}{n}$$

### 2) Median:

The median is the middle score for a set of data that has been arranged in order of magnitude.

In order to calculate the median, suppose we have the data below:

65, 55, 89, 56, 35, 14, 56, 55, 87, 45, 92

we first need to rearrange that data into order of magnitude.

14, 35, 45, 55, 55, 56, 65, 87, 89, 92

Median is 56.

### 3) Mode:

The mode is the most frequently occurring value in a set of values.

Eg: 14, 6, 11, 8, 7, 20, 11, 3, 7, 5, 7

7 occurs 3 times, Hence 7 is the mode.

## Variance:

It is the expectation of the squared deviation of a random variable from its mean.

If the generator of random variable X is discrete with probability mass function.

$x_1 \to p_1$ , $x_2 \to p_2$ , ... , $x_n \to p_n$

then
$$Var(x) = \sum_{i=1}^{n} p_i \cdot (x_i - \mu)^2 \text{ or}$$

$$Var(x) = \left(\sum_{i=1}^{n} p_i \cdot x_i^2\right) - \mu^2$$ where $\mu$ is the expected value.

i.e $\mu = \sum_{i=1}^{n} p_i x_i$ $\qquad Var = \frac{\sum (x_i - \mu)^2}{N}$

## Standard deviation:

It is a measure of the amount of variation or dispersion of a set of values.

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

## PROBABILITY:

Experiment: Any physical change (or process) which finally results into a definite outcome is simply known as an experiment.

Eg: Tossing a coin

Measuring the amount of rainfall in Banglore in the month of July.

Deterministic experiment: An experiment whose outcome can be determined in advance is called a deterministic experiment.

Eg: In a linear eqn $y = ax^2 + bx + c$, the values of a, b & c are fixed. So it is deterministic.

Random experiment: An experiment whose outcome can not be determined in advance, but is nevertheless still subjected to analysis is called as random experiment.

Outcome: The final result of an experiment is known as the outcome of the experiment.

Eg: While tossing a coin, the possible outcomes are head and ~~the~~ tail.

Sample space: The sample space of a random experiment is a mathematical abstraction used to represent all possible outcomes of the random experiment.

Eg: When a 6-faced dice is tossed, the sample space is,

$S = \{1, 2, 3, 4, 5, 6\}$

Finite Sample space: If the set of possible outcomes of the experiment is finite, then the associated sample space is a finite sample space.

Countably Infinite sample space: A sample space where the set of all outcomes can be put into a one-to-one correspondence with the natural number is said to be countably infinite.

Eg: Tossing a ~~fair~~ coin till a tail appears for the first time

$S = \{T, HT, HHT, HHHT, \ldots\}$

Event: Subsets of sample space are called events and can be represented by E, where $E \subseteq S$.

Eg: Suppose that a coin is tossed 3 times.

Sample space is $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

The event that the 3rd toss is tail is regardless of the first two tosses is

$E = \{HHT, HTT, THT, TTT\}$

Occurrence of an event: An event E of a random experiment is said to have occurred if the experiment terminates is an outcome that belongs to EA

Complement of an event: The complement of an event E, denoted by $\bar{E}$, is the event containing all points in S, but not in E.

Types of Events:

1) Exhaustive events:

Events are said to be exhaustive when they include all possibilities.

Eg: in tossing a coin either a head or tail comes up. There is no other possibility and therefore these are exhaustive events.

2) Favourable events: The elementary events which ensure the occurance of an event are called favourable events.

Eg: In tossing a 6-faced dice the number of an event are called favourable events to the occurance of a multiple of 3 are 2 i.e 3 and 6.

3) Mutually exclusive events:

Two events A and B are said to be mutually exclusive if they have no common elements.

i.e $A \cap B = \phi$. They are also called disjoint events.

If A & B are mutually exclusive, they cannot occur together.

Two events A and B are said to be mutually exclusive if the occurrence of one among them excludes the occurrence of the other.

i.e If A occurs, B does not occur and vice versa.

Ex: Suppose two 6 faced dice are tossed consecutively.

The Sample space is

$$S = \{(1,1), (1,2) \dots (1,6), (2,1), (2,2), \dots (2,6) \dots (6,1) \dots (6,6)\}$$

Let A be an event that the first dice is an odd number and the second dice is an even number.

$$A = \{(1,2)(1,4)(1,6)(3,2),(3,4),(3,6)(5,2)(5,4)(5,6)\}$$

Let B be an event that the first dice is a even number and the second dice is a odd number.

$$B = \{(2,1)(2,3)(2,5),(4,1)(4,3),(4,5),(6,1)(6,3)(6,5)\}$$

$A \cap B = \phi$

So, A & B are mutually exclusive events.

4) Null event or impossible event:

Any event that contains zero elements is a null event.

5) Simple or elementary event:

Any event that contains only one member of a sample space is called a simple event.

Ex: when tossing a 6-faced dice and let A be an event that face 3 is turned up, then S is a simple event.

6) Compound events: when an event is decomposable into a number of smaller events, then it is called a compound event.

For Ex: when tossing a two 6-faced dice, the event denoting the sum of two numbers of the two dice is 9 is a compound event as it can be decomposed into simple events like (3,6), (4,5), (5,4) & (6,1)

7) Equally likely events: Events are said to be equally likely if there is no reason to expect anyone in performace to any other. i.e when the probability of happening of two or more events is the same, they are called equally liked events.

8) De Morgan's laws: If A and B are two events then the complement of union of two events is same as the intersection of their complements and the complement of the intersection of two event is same as the union of their complements.

i) $\overline{A \cup B} = \overline{A} \cap \overline{B}$

ii) $\overline{A \cap B} = \overline{A} \cup \overline{B}$

PROBABILITY:

With each event $E_i$ in a finite sample space S, we associate a real number say $P(E_i)$ called the probability P of an event $E_i$ satisfying the following axioms:

1) $0 \leq P(E_i) \leq 1$

2) $P(S) = 1$ where the event S is the entire sample space called the "certain" event.

3) $P(E_1 \cup E_2 \cup \ldots \cup E_n) = P(E_1 + E_2 + E_3 + \ldots + E_n) = P(E_1) + P(E_2) + \ldots + P(E_n)$

or $P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i)$ where $E_1, E_2 \ldots E_n$ are n pairwise mutually exclusive events in S.

Addition rule of probability:

If two events A and B are mutually exclusive then,

$P(A \cup B) = P(A) + P(B)$

However, when A and B are not mutually exclusive i.e $A \cap B \neq \phi$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ex.1) A bag contains 18 colored marbels. Out of which 4 are colored red, 8 are colored yellow & 6 are colored green. A marbel is selected at random. What is the probability that the ball chosen is either red or green?

Let A be an event of choosing a red marble & B be an event of choosing a green marble.

$P(A) = \frac{4}{18}$ $P(B) = \frac{6}{18}$

$P(A \cap B) = 0$ (no ball can be simultaneously red & green)

By Addition rule of probability.

$P(A \cup B) = P(A) + P(B) = \frac{4}{18} + \frac{6}{18} = \frac{10}{18} = \frac{5}{9}$

22] Consider a pack of 52 play. cards. A card is selected at random. what is the probability that the card is either a diamond or king.

Let A be an event of choosing diamond card & B be an event of choosing king card.

$P(A) = \frac{13}{52}$ $P(B) = \frac{4}{52}$ $P(A \cap B) = \frac{1}{52}$

$P(A \cup B) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}$

**Joint Probability:**

The probability of occurrence of both events A and B together denoted by $P(A \cap B)$ is known as the Joint probability of A & B.

**Marginal probability:**

A probability of only one event A that takes place irrespective of any other event is called a marginal probability

6 balls, there 3 are red, 3 are blue. picking a red is

$P(R) = 3/6 = 1/2$

In other words $P(B|A)$ is the no. of outcomes in $A \cap B$ divided by the number of outcomes in A.

$$P(B|A) = \frac{\text{no. of outcomes in } A \cap B}{\text{no. of outcomes in } A}$$

**Conditional probability:**

The conditional probability of event B given A has occurred is defined as:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

if $P(A) \neq 0$

2.3) Suppose that a fair dice is tossed twice. Events A & B are defined as follows:

$A = \{$ the first dice shows an even number $\}$

$B = \{$ the second dice shows either 5 or 6 $\}$

Compute the following:

1) $P(A|B)$ 2) $P(B|A)$

Sol: $S = \{(1,1)(1,2)\ldots(1,6)(2,1)\ldots(2,6)(3,1)\ldots(3,6)(4,1)\ldots(4,6)(5,1)\ldots(5,6)(6,1)\ldots(6,6)\}$

$A = \{(2,1)(2,2)\ldots(2,6),(4,1)(4,2)\ldots(4,6),(6,1)\ldots(6,6)\}$

$B = \{(1,5)(1,6)(2,5)(2,6)(3,5)(3,6)(4,5)(4,6)(5,5)(5,6)(6,5)(6,6)\}$

$P(A) = \frac{18}{36}$ $P(B) = \frac{12}{36}$ $P(A \cap B) = \frac{6}{36}$

1) $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{6/36}{12/36} = \frac{1}{2}$

2) $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{6/36}{18/36} = \frac{1}{3}$

2.5) A fair coin is tossed three times successively. Find the conditional probability. $P(A|B)$, where the events A & B are as follows:

A = "more heads than tails come up"

B = "getting tail in the first toss".

$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

$P(A) = \frac{4}{8}$ $P(B) = \frac{4}{8}$ $P(A \cap B) = \frac{1}{8}$

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/8}{4/8} = \frac{1}{4}$

## Independent events:

If the occurrence of one event A does not effect, nor is affected by the occurrence of another event B, then we say that A & B are independent events.

If A & B are independent, then

$P(B|A) = P(B)$ $P(A|B) = P(A)$

Ex: Picking a ball from a box & getting heads after tossing a coin.

Getting a 2 on a 6 faced die in the first roll & then getting a 6 on a second roll of the dice

## Multiplication rule:

The multiplication rule is defined below.

The events A and B are said to be independent if

$$P(A \cap B) = P(A) \cdot P(B)$$

In other words, the probability of independ events A and B occuring is the product of the probabilities of the events occuring seperately.

Suppose, if the event A & B are not independent, then

$$P(A \cap B) = \begin{cases} P(B) \cdot P(A|B) & \text{if } P(B) \neq 0 \\ P(A) \cdot P(B|A) & \text{if } P(A) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

2.6] Suppose that five good fuses and two defective ones have been mixed up. To find the defective fuses, we test them one-by-one at random and without replacement. What is the probability that we are lucky and find both of the defective fuses in the first two tests?

$P(A \cap B) = ?$

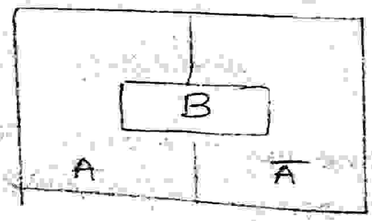
$P(A)$ = Probability of getting defective fuse in the first test & $P(B)$ in $2^{nd}$ test.

$P(A) = \frac{2}{7}$ $\quad P(B) = \frac{1}{6}$

$P(A \cap B) = \frac{2}{7} \cdot \frac{1}{6} = \frac{2}{42}$

## Law of Total Probability:

It is a fundamental rule relating marginal probabilities to to conditional probabilities.



$P(B) = P(A \cap B) \cup P(\bar{A} \cap B)$

$= P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})$

by the definition of conditional probability & multiplication rule of probability.

Let $A_1, A_2, \ldots, A_n$ be n mutually disjoint events that form the partition of the sample space S & assume that $P(A_i) > 0$ for all i. Then for any event B

$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B).$$

$$P(B) = P(A_1 \cap B) \cup P(A_2 \cap B) \cup \ldots \cup P(A_n \cap B)$$
$$= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + \ldots + P(B|A_n) \cdot P(A_n)$$
$$= \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)$$

This relation is called as the law of total probability. and is sometimes called the rule of elimination.

2.8) Consider 3 boxes $B_1$, $B_2$ and $B_3$ consisting of 10 balls each. $B_1$ contains 4 red balls and 6 blue balls, $B_2$ contains 7 red balls and 3 blue balls and box $B_3$ contains 5 red balls and 5 blue balls. Suppose out of 3 boxes one box is chosen at random and a ball is drawn from it at random. What is the probability that the chosen ball is red?

Let $B_i$ be an event that the box $i$ is chosen & R be an event that the chosen ball is red.

$P(R|B_1) = \frac{4}{10} = 0.4$ $\quad P(R|B_2) = \frac{7}{10} = 0.7$ $\quad P(R|B_3) = \frac{5}{10} = 0.5$

$P(B_1) = 1/3$ $\quad P(B_2) = 1/3$ $\quad P(B_3) = 1/3$

$$P(R) = P(R \cap B_1) \cup P(R \cap B_2) \cup P(R \cap B_3)$$
$$= P(R|B_1) \cdot P(B_1) + P(R|B_2) \cdot P(B_2) + P(R|B_3) \cdot P(B_3)$$
$$= 0.4 \times \frac{1}{3} + 0.7 \times \frac{1}{3} + 0.5 \times \frac{1}{3}$$
$$= \underline{0.5333}$$

2.9) The probability of winning a cricket tournament by India against Pastkin is 0.4, 0.6 against Srilanka and 0.8 against Bangladesh. India plays a game against a randomly chosen opponent. What is the probability of winning?

Let A, B and C be the events of playing against Pakistan Srilanka and Bangladesh resp.

Let W be the probability of winning.

Then $P(A) = 1/3$ $\quad P(B) = 1/3$ $\quad P(C) = 1/3$

$P(W|A) = 0.4$ $\quad P(W|B) = 0.6$ $\quad P(W|C) = 0.8$

The law of total probability, the probability of winning is:

$$P(W) = P(W \cap A) \cup P(W \cap B) \cup P(W \cap C)$$
$$= P(W|A) \cdot P(A) + P(W|B) \cdot P(B) + P(W|C) \cdot P(C)$$
$$= 0.4 \times \frac{1}{3} + 0.6 \times \frac{1}{3} + 0.8 \times \frac{1}{3}$$
$$= \underline{0.6}$$

Bayes theorem:

Given a hypothesis H and evidence E, Bayes theorem states that the relationship between the probability of the hypothesis before getting the evidence P(H) & the probability of the hypothesis after getting the evidence P(H|E) is:

$$P(H/E) = \frac{P(E|H) \cdot P(H)}{P(E)}.$$

A prior probability: is an initial probability value originally obtained before any additional information ~~are~~ ~~commonly used~~ is obtained.

A posterior probability: is a probability value that has been revised by using additional information that is later obtained.

For a given partition of S into n mutually exclusive sets $A_1, A_2 - A_n$, we want to know that probability that some particular case $A_i$, occurs given that some event B, occurs.

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}$$

Using multiplication rule,

$$P(A_i \cap B) = P(B|A_i) P(A_i)$$

Using the law of total probability,

$$P(B) = \sum_{i=1}^{n} P(B/A_i) \cdot P(A_i)$$

$$\therefore \boxed{P(A_i|B) = \frac{P(B/A_i) \cdot P(A_i)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}}$$

This eqn is called the Bayes' rule.

When to use Bayes' theorem?

1) The sample space is partitioned into a set of mutually exclusive events $A_1, A_2 \dots A_n$.

2) Within the sample space there exists an event B for which $P(B) > 0$

3) The analytical goal is to compute a conditional probability of the form: $P(A_i|B)$

4) You know at least one of the two sets of probabilities described below

1. $P(A_i \cap B)$ for each A

2.10] Let A be an event of choosing a king Card. Then $P(A) = \frac{4}{52} = \frac{1}{13}$.
In other words, the prior probability is $P(A) = \frac{1}{13}$. Suppose an evidence is provided that the card is a face card, call it event B. Then posterior probability $P(A|B)$ can be calculated using Bayes' theorem as:

$$P(A/B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{1 \cdot 1/13}{3/13} = \frac{1}{3}$$

Since every king is also a face then $P(B/A) = 1$
Since there are 3 face cards, then $P(B) = \frac{12}{52} = \frac{3}{13}$
The likelihood ratio is $\frac{1}{3/13} = \frac{13}{3}$
∴ the probability that the randomly chosen card is a king given the evidence that the card chosen is a face card is $1/3$.

2.12] A box contains 10 red and 12 blue ~~bl~~ balls. Two balls are drawn at random and are discarded without their colors being seen. What is the probability that a third ball drawn is blue?

Let $R_1$ be an event that the $1^{st}$ ball drawn is Red.
$B_1$ —"— Blue
$R_2$ —"— red
$B_2$ —"— blue.

Let $R_1 B_2$ denote the first ball drawn is a red and the second ball drawn is blue.

$$P(R_1 R_2) = \frac{{}^{10}C_1}{{}^{22}C_1} \times \frac{{}^{9}C_1}{{}^{21}C_1} = \frac{10}{22} \times \frac{9}{21}$$

$$P(B_1 B_2) = \frac{{}^{12}C_1}{{}^{22}C_1} \times \frac{{}^{11}C_1}{{}^{21}C_1} = \frac{12}{22} \times \frac{11}{21}$$

$$P(B_1 R_2) = \frac{{}^{12}C_1}{{}^{22}C_1} \times \frac{{}^{10}C_1}{{}^{21}C_1} = \frac{12}{22} \times \frac{10}{21}$$

$$P(R_1 B_2) = \frac{{}^{10}C_1}{{}^{22}C_1} \times \frac{{}^{12}C_1}{{}^{21}C_1} = \frac{10}{22} \times \frac{12}{21}$$

Let $P(E)$ be the probability of choosing $3^{rd}$ ball is blue.
1st two balls are $R_1 B_2 \Rightarrow P(E_1) = \frac{{}^{11}C_1}{{}^{20}C_1} = \frac{11}{20}$

$R_1 R_2 \Rightarrow P(E_2) = \frac{{}^{12}C_1}{{}^{20}C_1} = \frac{12}{20}$

$B_1 B_2 \Rightarrow P(E_3) = \frac{{}^{10}C_1}{{}^{20}C_1} = \frac{10}{20}$

$B_1 R_2 \Rightarrow P(E_4) = \frac{{}^{11}C_1}{{}^{20}C_1} = \frac{11}{20}$

$P(B_3) = P(B_3|R_1B_2) \cdot P(R_1B_2) + P(B_3|R_1R_2) \cdot P(R_1R_2) +$
$P(B_3|B_1R_2) \cdot P(B_1R_2) + P(B_3|B_1B_2) \cdot P(B_1B_2)$

$P(R_1R_2B_3) = \frac{10}{22} \times \frac{9}{21} \times \frac{12}{20}$

$P(R_1B_2B_3) = \frac{10}{22} \times \frac{12}{21} \times \frac{11}{20}$

$P(B_1B_2B_3) = \frac{12}{22} \times \frac{11}{21} \times \frac{10}{20}$

$P(B_1R_2B_3) = \frac{12}{22} \times \frac{10}{21} \times \frac{11}{20}$

$P(B_3) = \frac{10}{22} \times \frac{9}{21} \times \frac{12}{20} + \frac{10}{22} \times \frac{12}{21} \times \frac{11}{20} + \frac{12}{22} \times \frac{11}{21} \times \frac{10}{20} +$
$\frac{12}{22} \times \frac{10}{21} \times \frac{11}{20}$

$= \frac{1080 + 1320 + 1320 + 1320}{9240}$

$= \frac{5040}{9240} = 0.5454.$

---

## UNIT - 2

### RANDOM VARIABLES AND PROBABILITY DISTRIBUTION

A random variable is a numerical description of the outcome of a statistical experiment.

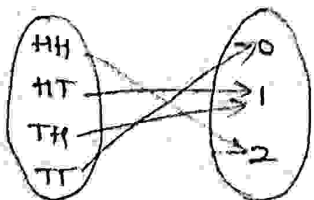
X is a random variable.

Suppose 2 coins are tossed simultaneously.

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

Let X denotes the counting no. of heads.

X can take the values 0, 1, 2




Domain of X      Range of X

$P(X=0) = P(0) = 1/4$ for the event $x=0$ equivalent to event {TT}

$P(X=1) = P(1) = 1/2$ —"— $x=1$ —"— {TH, HT}

$P(X=2) = P(2) = 1/4$. —"— $x=2$ —"— {HH}

## Probability Distribution of $x$:

The mapping of all the possible values of a random variable $x$ to their corresponding probabilities is known as the probability distribution of $x$.

It is also denoted as $f(x)$.

1) A discrete random variable has a countable number of possible values.

2) Continuous random variable

The probability of each value of a discrete random variable is b/w 0 & 1 & sum of all the probabilities is equal to 1

| $x$ | $x_1$ | $x_2$ | ... | $x_n$ |
|---|---|---|---|---|
| $P(X=x)$ | $P(X=x_1)$ | $P(X=x_2)$ | ... | $P(X=x_n)$ |

Probability distribution or probability function of the random variable $X$.

Consider the experiment ~~of~~ of tossing 2 coins

3 possible outcomes and their probabilities are

1) No heads: $X(0)=\{TT\}$; $P(X=0)=\frac{1}{4}$

2) One head: $X(1)=\{TH, HT\}$; $P(X=1)=\frac{1}{2}$

3) Two heads: $X(2)=\{HH\}$; $P(X=2)=\frac{1}{4}$.

The probability distribution is

| $X$ | 0 | 1 | 2 |
|---|---|---|---|
| $P(X=x)$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

## Probability mass Function (pmf):

Let $f_X(x)$ be a function for a discrete random variable $X$, then function, $f(x): P(x)=P(X=x)$ is called probability mass function of $X$

Properties of the probability distribution for a discrete random variable:

A function $f_X(x)$ is a probability mass function for discrete random variable $X$ with random $R_X$ of the form $\{x_1, x_2, ...\}$ if and only if it satisfy the following two conditions:

1) $f_X(x) \geq 0$ for each value within its domain

2) $\sum f_X(x)=1$

Consider an experiment of tossing 3 coins simultaneously. X be an random variable denoting no. of heads give the probability distribution of $x$ in terms of probability distribution table and venn diagram.

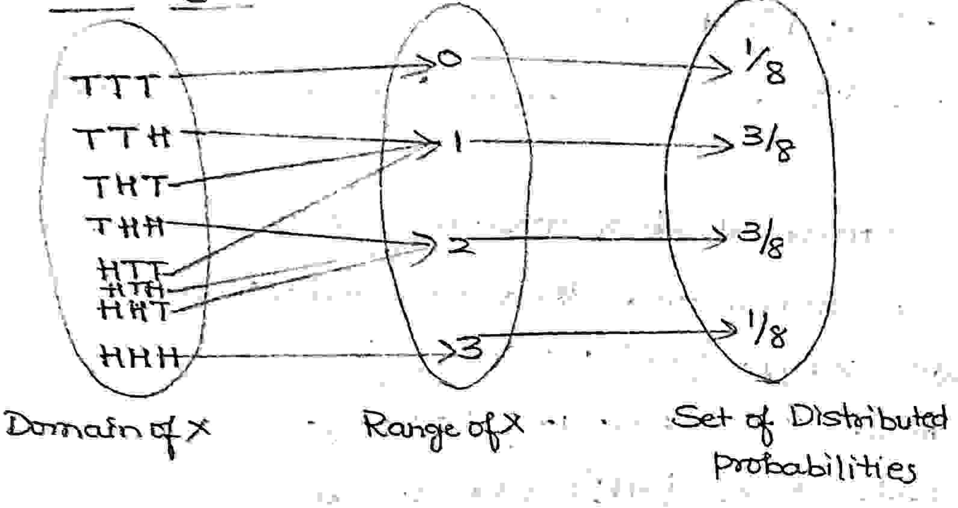
$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

X can take the value 0, 1, 2, 3

Probability distribution table

| $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X=x)$ | $1/8$ | $3/8$ | $3/8$ | $1/8$ |

Venn diagram:




3.2] Consider an experiment when a pair of two dice is tossed. Let X and Y be two random variables on the sample spaces where X = maximum of 2 numbers i.e $X(a,b) = max(a,b)$ & Y = sum of two numbers i.e $Y(a,b) = a+b$. Find

1) probability distribution function $f_X(x)$ and
2) probability distribution function $g_Y(y)$
3) The probability distribution graph for $f_X(x)$ & $g_Y(y)$.

$S = \{(1,1)(1,2)(1,3)(1,4)(1,5)(1,6)(2,1)(2,2)(2,3)(2,4)(2,5)(2,6)(3,1)$
$(3,2)(3,3)(3,4)(3,5)(3,6)(4,1)(4,2)(4,3)(4,4)(4,5)(4,6)$
$(5,1)(5,2)(5,3)(5,4)(5,5)(5,6)(6,1)(6,2)(6,3)(6,4)(6,5)(6,6)\}$

1) X can take the values 1, 2, 3, 4, 5, 6

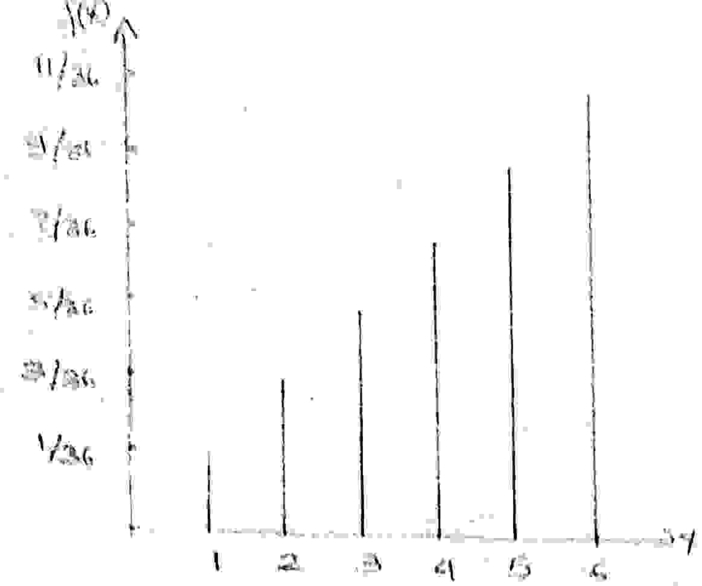
$f(1) = 1/36$ $f(2) = 3/36$ $f(3) = 5/36$ $f(4) = 4/36$ $f(5) = 9/36$

$f(6) = 1/36$

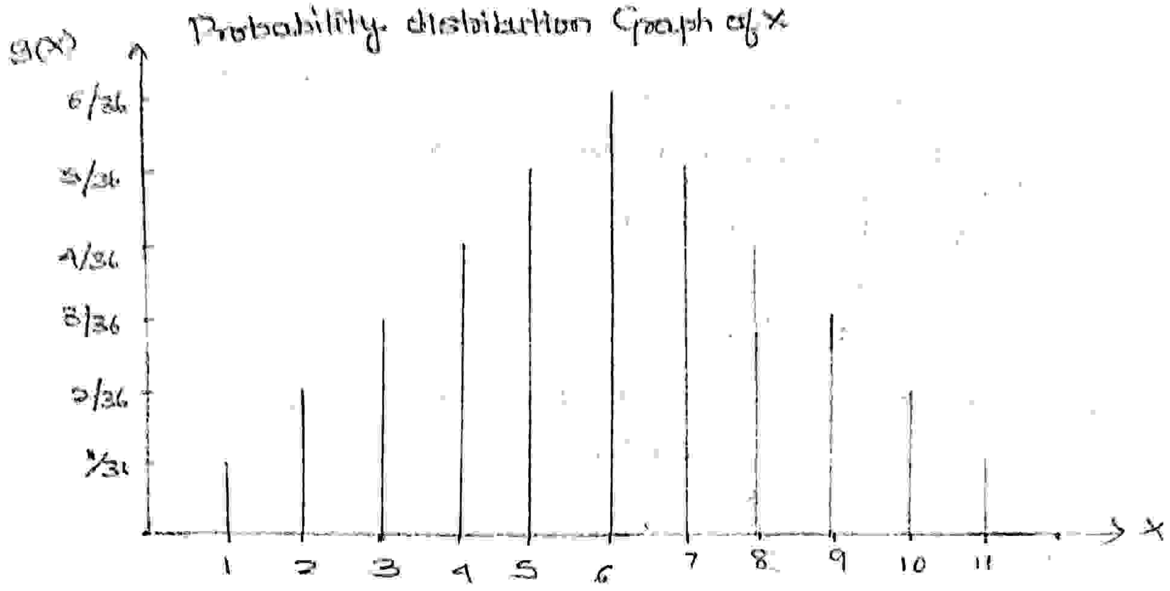
2) Y can take the values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

$g(2) = 1/36$ $g(3) = 2/36$ $g(4) = 3/36$ $g(5) = 4/36$ $g(6) = 5/36$

$g(7) = 6/36$ $g(8) = 5/36$ $g(9) = 4/36$ $g(10) = 3/36$ $g(11) = 2/36$ $g(12) = 1/36$




Probability distribution Graph of X




Probability Distribution Graph of Y

Cummulative distribution function (CDF):

The cummulative distribution function (cdf) of a discrete random variable X is the function F(t) which tells you the probability that X is less than or equal to t.

Properties of a CDF:

1> $F(x)$ always lies between 0 & 1.

2> $F(x)$ as $x \to -\infty = 0$

3> $F(x)$ as $x \to \infty = 1$

4> Suppose $x_1 < x_2$ then $F(x_1) < F(x_2)$.

Q Consider a coin tossing ~~as~~ 3 times. Let $x$ be a random variable denoting the no. of heads, that come. Determine cummulative distribution of function.

$S = \{HHH, HTH, HHT, HTT, THH, THT, TTH, TTT\}$

$f(x=0) = P(X=0) = 1/8$

$f(1) = P(X=1) = 3/8$

$f(2) = P(X=2) = 3/8$

$f(3) = P(X=3) = 1/8$

Cummulative distribution of function are

$F(0) = \frac{1}{8}$

$F(1) = f(0) + f(1) = \frac{1}{2}$

$F(2) = f(0) + f(1) + f(2) = 7/8$

$F(3) = f(0) + f(1) + f(2) + f(3) = 1$

$$F_X(x) = \begin{cases} 0 & x<0 \\ 1/8 & 0\le x<1 \\ 4/8 & 1\le x<2 \\ 7/8 & 2\le x<3 \\ 1 & x\ge 3 \end{cases}$$

2. A random variable x has following probability distribution function

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(x)$ | $k$ | $3k$ | $5k$ | $7k$ | $9k$ |

Find i) the value of k

ii) a) $P(x<3)$ b) $P(x\ge 3)$ and c) $P(0<x<4)$

iii) the distribution function of x.

i) $\sum_{i=1}^{n} P(x_i) = 1$

$k + 3k + 5k + 7k + 9k = 1$

$25k = 1$

$k = 1/25$

ii) a) $P(X<3) = P(X=0) + P(X=1) + P(X=2)$

$= k + 3k + 5k$

$= 9k$

$= \frac{9}{25}$

b) $P(X\ge 3) = P(X=3) + P(X=4)$

$= 7k + 9k$

$= 16/25$

c) $P(0<x<4) = 3k + 5k + 7k = \frac{15}{25}$

iii) $F(0) = P(x=0) = \frac{1}{25}$

$F(1) = P(X=0) + P(X=1) = k + 3k = 4/25$

$F(2) = P(X=0) + P(X=1) + P(X=2) = k + 3k + 5k = 9/25$

$F(3) = k + 3k + 5k + 7k = 16/25$

$F(4) = k + 3k + 5k + 7k + 9k = \frac{25}{25}$

3. A random variable has the following probability distribution function:

| X | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| $P(X=x)$ | 0 | $k$ | $2k$ | $2k$ | $3k$ | $k^2$ | $2k^2$ | $7k^2+k$ |

Find i) the value of k

ii) Evaluate $P(X<6)$

iii) Evaluate $P(X\ge 6)$

iv) If $P(X\le c) > \frac{1}{2}$, then find the minimum value of c

i) $\sum_{i=1}^{n} P(x_i)=1$

$0+k+2k+2k+3k+k^2+2k^2+7k^2+k=1$

$10k^2+9k-1=0$

$10k^2+10k-k-1=0$

$10k(k+1)-1(k+1)=0$

$10k-1=0$

$10k=1$

$k=1/10$

ii) $P(x<6)=0+k+2k+2k+3k+k^2$

$=k^2+8k$

$=\frac{1}{100}+\frac{8}{10}$

$=\frac{81}{100}$

iii) $P(x\geq 6)=2k^2+7k^2+k$

$=9k^2+k$

$=\frac{9}{100}+\frac{1}{10}$

$=\frac{19}{100}$

4 iv) Compute the value of $F(x)$ until $P(X\leq c)>\frac{1}{2}$

$F(0)=P(X=0)=0$

$F(1)=P(X=0)+P(X=1)=0+k=1/10$

$F(2)=P(X=0)+P(X=1)+P(X=2)=3k=3/10$

$F(3)=k+2k+2k=5k=5/10=1/2$

$F(4)=k+2k+2k+3k=8k=8/10=\frac{4}{5}$

~~$F(5)=k+2k+2k+3k+k^2=8k+k^2=$~~

$\therefore$ Minimum value of c is 4.

3.7) Let X be a discrete random variable whose distribution function is given by:

$$F(x)=\begin{cases} 0 & x<-3 \\ 1/6 & -3\leq x<6 \\ 1/2 & 6\leq x<10 \\ 1 & 10\leq x \end{cases}$$

i) Find $P(X\leq 4)$

ii) $P(-5<x\leq 4)$

iii) The probability distribution of x

i) $P(X\leq 4)=F(4)=1/6$

ii) $P(-5<X\leq 4)=F(4)-F(-5)=P(X\leq 4)-P(X\leq -5)$

$=1/6-0=1/6$

3. The probability distribution of X is given in the following table

| $x$ | $-3$ | $6$ | $10$ |
|---|---|---|---|
| $P(X=x)$ | $1/6$ | $2/6$ | $1/2$ |

Probability of a continuous random variable:

The probability of a continuous random variable X is defined as the probability of X that falls in some interval i.e $P(a<X<b)$, where a and b are constants.

Then for any two constants a and b with $a<b$,

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

For any pre-determined value $x$, $P(X=x)=0$, since if we measure X accurately enough, we are never going to hit the value $x$ exactly.

Probability Density Function (pdf):

simply density function:

Let X be a random variable. There exists a function $f(x)$ so that for any constants a and b, with $-\infty \le a \le b \le \infty$

$$P(a \le X \le b) = \int_a^b f(x)dx$$

The function $f(x)$ is called the probability density function.

For any a,

$$P(X=a) = P(a \le X \le a) = \int_a^a f(x)dx$$

Properties:

1) Probability Density function always lies on or above the x-axis.

2) The area under the density curve is 1.

Difference between pmf and pdf:

The probability density function $f(x)$ of a continuous random variable is the analogue of the probability mass function $p(x)$ of a discrete random variable.

Two important differences:

1) Unlike $p(x)$, the pdf $f(x)$ is not a probability. You have to integrate it to get probability.

2) Since $f(x)$ is not a probability there is no restriction that $f(x)$ be less than or equal to 1.

Cumulative Distribution Function for a Continuous Random Variable:

The CDF of a continuous random variable is defined exactly the same as for discrete random variables.

If $X$ is a random variable with probability density function $f(x)$, then the associated cummulative distribution function is defined by:

$$F(a) = P(X \leq a) = P(X \in (-\infty, a]) = \int_{-\infty}^{a} f(x)\,dx.$$

The CDF is found by integrating the pdf between the minimum value of $X$ and $a$.

The pdf of a continuous random variable can be obtained by differenting the CDF.

$$\frac{dF}{dx} = f(x)$$

The CDF can be used to find the probability of a random variable being between two values.

$P(a \leq X \leq b)$ = the probability that $X$ is between $a$ and $b$

But this equals to the probability that $X \leq a$ minus the probability that $X \leq b$.

$$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

$F(x)$ has the following properties:

1) $0 \leq F(x) \leq 1$, $-\infty < x < \infty$

2) $F$ is monotonically increasing function i.e $F(a) \leq F(b)$ whenever $a \leq b$.

3) The limit of $F$ to the left is 0 and to the right is 1

i.e $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$

4) The probability associated with the event $[x = c]$ is zero.

therefore:

$$P(X = c) \quad P(c \leq X \leq c) = \int_{c}^{c} f_X(x)\,dx = 0$$

3.8] The function $f(x) = \begin{cases} e^{-x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$ is a probability density for the random variable $X$. Compute $P(-10 \leq X \leq 10)$

$$P(-10 \leq X \leq 10) = \int_{-10}^{10} f(x)\,dx = \int_{-10}^{0} 0\,dx + \int_{0}^{10} e^{-x}\,dx$$

$$= 0 + \frac{e^{-10}}{-1} - \frac{e^{-0}}{-1} = 1 - e^{-10}$$

3.9] Suppose the income of people in a community can be approximated by a continuous probability distribution with density $f(x) = \begin{cases} 2x^{-2} & \text{if } x \geq 2 \\ 0 & \text{if } x < 0 \end{cases}$

Find the probability that a randomly chosen person has an income

1) between \$30,000 to \$50000

2) of at least \$60000

3) of at most \$40000

Let X be the income of a randomly chosen person.

1) The probability that a randomly chosen person has an income between \$30,000 and \$50000

$$P(3 \leq X \leq 5) = \int_3^5 2x^{-2}\,dx$$

$$= \frac{2x^{-2+1}}{-2+1}\Big|_3^5 = -2x^{-1}\Big|_3^5 = \frac{-2}{x}\Big|_3^5$$

$$= \frac{-2}{5} + \frac{2}{3}$$

$$= \frac{-6+10}{15} = \frac{4}{15}$$

2) $$P(X \geq 6) = \int_6^\infty 2x^{-2}\,dx = \frac{-2}{x}\Big|_6^\infty$$

$$= \frac{-2}{\infty} + \frac{2}{6}$$

$$= \frac{1}{3}$$

3) $$P(X \leq 4) = \int_{-\infty}^{2} 0\,dx + \int_2^4 2x^{-2}\,dx$$

$$= \frac{-2}{x}\Big|_2^4 = \frac{-2}{4} + \frac{2}{2}$$

$$= -\frac{1}{2} + 1$$

$$= 1/2$$

3.10] The amount of time in hours that an electric bulb functions before breaking down is a continuous random variable with pdf given by:

$$f(x) = \begin{cases} xe^{-x/100} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

What is the probability that 1) The bulb will function b/w 200 to 300 hrs before breaking down and
2) It will function for less than 250 hrs?

1) WKT, $\int_{-\infty}^{\infty} f(x)\,dx = 1$,

Then $\lambda \int_{0}^{\infty} e^{-\frac{x}{100}}\,dx = 1$

$$\lambda \left. \frac{e^{-x/100}}{-1/100} \right|_0^{\infty} = 1$$

$$-\lambda 100 \left. e^{-x/100} \right|_0^{\infty} = 1$$

$$-\lambda 100 \left( e^{-\infty/100} - e^{0/100} \right) = 1$$

$$-\lambda 100 \left[ e^{-\infty} - e^{0} \right] = 1$$

$$-\lambda 100 \left( -e^{0} \right) = 1$$

$$\lambda 100 = 1$$

$$\lambda = \frac{1}{100}$$

2) $\int_{-\infty}^{0} 0\,dx + \int_{0}^{250} \lambda e^{-x/100}\,dx$

$$= \lambda \left. \frac{e^{-x/100}}{-1/100} \right|_0^{250} = -100\lambda \left( e^{-\frac{250}{100}} - e^{-0/100} \right)$$

$$= -100 \times \frac{1}{100} \left( e^{-25/10} - e^{0} \right)$$

$$= 1 - e^{-25/10} = 1 - 0.08208$$

$$= 0.9179$$

EXPECTATION

The distribution function $F(x)$ or the density $f(x)$ [pmf $p(x_i)$ for a discrete random variable] completely characterized by the behaviour of a random variable $X$.

The mean, median and mode are often called measure of central tendency of a random variable $X$.

Definition (Expectation): The expectation, $E[X]$ of a random variable $X$ is defined by

$$E[X] = \begin{cases} \sum_i x_i\, p(x_i), & \text{if } X \text{ is discrete.} \\ \int_{-\infty}^{\infty} x f(x)\,dx, & \text{if } X \text{ is continuous.} \end{cases}$$

The relavant sum or integral is absolutely convergent. that is $\sum_i x_i\, p(x_i) < \infty$ and $\int_{-\infty}^{\infty} x f(x)\,dx < \infty$

Variance:

The variance of a random variable X is

$$Var[X] = \mu_2 = \sigma^2_x = \begin{cases} \sum_i (x_i - E[X])^2 p(x_i) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} (x - E[X])^2 f(x)dx & \text{if } X \text{ is continuous} \end{cases}$$

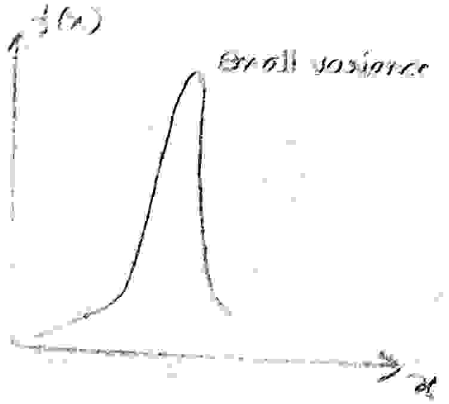
Var[X] is always a nonnegative number.

Standard deviation:

The square root $\sigma_x$ of the variance is known as the standard deviation

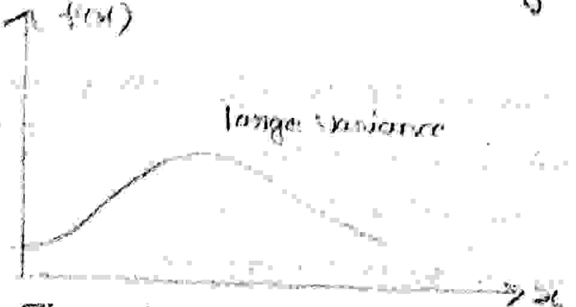
The variation and the standard deviation are measures of the 'spread' or 'dispersion' of a distribution.

If X has a concentrated distribution so that X takes values near to E[X] with a large probability, then the variance is small




The pdf of a concentrated distribution

Diffuse distribution - one with a large value of $\sigma^2$.




The pdf of a diffuse distribution

The no. of failures of a computer system in a week of operation has the following pdf.

| No. of failures | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Probability | 0.18 | 0.28 | 0.25 | 0.18 | 0.06 | 0.04 | 0.01 |

a) Find the expected number of failures in a week

b) Find the variance of the number of failures in a week.

a) $E[X] = \sum_i x_i P(x_i)$

$= 0\,(0.18) + 1\,(0.28) + 2\,(0.25) + 3\,(0.18) + 4\,(0.06) + 5\,(0.04) + 6\,(0.01)$

$= \underline{1.82}$

b) $Var[X] = \sigma_x^2 = \sum_i (x_i - E[X])^2 P(x_i)$

$= (0-1.82)^2(0.18) + (1-1.82)^2(0.28) + (2-1.82)^2(0.25) + (3-1.82)^2(0.18) + (4-1.82)^2(0.06) + (5-1.82)^2(0.04) + (6-1.82)^2(0.01)$

$= \underline{1.9076}$

## EXPECTATION BASED ON MULTIPLE RANDOM VARIABLES

Let $X_1, X_2, \dots, X_n$ be $n$ random variables defined on the same probability space and let $Y = \phi(X_1, X_2, \dots X_n)$.

Then $E[Y] = E[\phi(x_1, x_2, \dots, x_n)]$

$$= \begin{cases} \sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} \phi(x_1, x_2, \dots x_n)\, P(x_1, x_2, \dots x_n) & \text{discrete} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \phi(x_1, x_2, \dots x_n)\, f(x_1, x_2, \dots x_n)\, dx_1 dx_2 \dots dx_n & \text{continuous} \end{cases}$$

**THEOREM:** (The Linearity Property of Expectation)

Let $X$ and $Y$ be two random variables. Then the expectation of their sum is the sum of their expectations.

that is, if $Z = X+Y$, then $E[Z] = E[X+Y] = E[X] + E[Y]$

Proof: We will prove the theorem assuming that $X$, $Y$ and hence $Z$ are continuous random variables.

The proof for the discrete case is very similar.

$$E[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x,y)\, dy\, dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx + \int_{-\infty}^{\infty} y f_Y(y)\, dy \quad \text{[By defn of the marginal densities]}$$

$$= E[X] + E[Y].$$

This theorem does not require that $X$ and $Y$ be independent.

It can be generalised to the case of $n$ variables.

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

and to $E\left[a_0 + \sum_{i=1}^{n} a_i X_i\right] = a_0 + \sum_{i=1}^{n} a_i E[X_i]$.

where $a_0, a_1 \ldots a_n$ are constants.

For instance, let $X_1, X_2, X_3 \ldots X_n$ be random variables (not necessarily independent) with a common mean $\mu = E[X_i]$ $(i = 1, 2, \ldots, n)$

Then the expected value of their sample mean is equal to $\mu$.

$$E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \mu$$

Standard distribution:

I) Descrete distribution:

1) Binomial distribution
2) Hyper geometric distibution
3) Poisson's distribution
4) Geometric distribution
5) Negative binomial distribution

II) Continuous distribution:

1) Uniform or rectangular distribution
2) Exponential distribution
3) Gamma distribution
4) Weibul distribution
5) Normal distribution
6) Logistic
7) Beta

Properties of binomial experiment:

1) There must be a fixed number of trials
2) All trials should have identical properties.
3) The probability of success is constant in each trials.
4) All trials must be independent of each other.

Pmf of binomial distribution:

Consider a set of $n$ independent trials.
Let $p$ denote the probability of success in a trial.
$q = 1-p$ denote the probab
Then a random variable $X$ is said to follow binomial distribution if it takes on only non negative values and its probability mass function is given by:

$$P(X=k) = P(k) = {}^{n}C_{k}\, p^{k} q^{n-k} \quad ; k = 0, 1, 2, \ldots n$$

Mean and variance of the binomial distribution:

Any random variable with a binomial distribution X with parameters n and p is a sum of n independent Bernouli random variables in which the probability of success is p.

$X = X_1 + X_2 + X_3 + \cdots + X_n$

The mean and variance of each $X_i$ can easily be calculated as:

$E(X_i) = P.$

$Var(X_i) = P(1-P) = nq$

Hence, the mean and variance of X are given by

$\mu = E(X) = np$

$\sigma^2 = Var(X) = np(1-P) = npq$

A biased coin is tossed 6 times. The probability of getting heads on any toss is 0.3. Let X denote the no of heads that come up.

Calculate: a) $P(X=2)$

b) $P(X=3)$

c) $P(1 < X \leq 5)$

$n=6$

a) $P(X=2) = {}^6C_2 P^2 q^{6-2}$

$= \frac{6 \times 5}{2} \cdot (0.3)^2 (0.7)^4 = \underline{\underline{0.3241}}$

b) $P(X=3) = {}^6C_3 P^3 q^{6-3}$

$= \frac{6 \times 5 \times 4}{3 \times 2} (0.3)^3 (0.7)^3$

$= \underline{\underline{0.18522}}$

c) $P(1 < X \leq 5) = P(X=2) + P(X=3) + P(X=4) + P(X=5)$

$= 0.3241 + 0.18522 + {}^6C_4 (0.3)^4 (0.7)^2 + {}^6C_5 (0.3)^5 (0.7)$

$= \underline{\underline{0.5790}}$

2) The probability that sachine hits a target at any time is $p = 1/4$. Suppose he fires at the target 7 times. Find the probability that he hits the target

(a) Exactly 3 times

(b) At least 1 time

(c) Find expectation, variance and standard deviation.

$n = 7$, $p = \frac{1}{4}$, $q = 1 - \frac{1}{4} = 3/4$

a) Exactly 3 times

$$P(X=3) = {}^7C_3\, p^3 q^{7-3}$$

$$= \frac{7\times6\times5}{3\times2} (1/4)^3 (3/4)^4$$

$$= 0.1730$$

b) At least one time

$$P(X \geq 1) = P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7)$$

$$= {}^7C_1 (1/4)^1 (3/4)^6 + {}^7C_2 (1/4)^2 (3/4)^5 + {}^7C_3 (1/4)^3 (3/4)^4 + {}^7C_4 (1/4)^4 (3/4)^3$$
$$+ {}^7C_5 (1/4)^5 (3/4)^2 + {}^7C_6 (1/4)^6 (3/4)^1 + {}^7C_7 (1/4)^7 (3/4)^0$$

$$= 7(0.25)(0.75)^6 + \frac{7\times6}{2}(0.25)^2(0.75)^5 + \frac{7\times6\times5}{3\times2}(0.25)^3(0.75)^4 +$$
$$\frac{7\times6\times5}{3\times2}(0.25)^4(0.75)^3 + \frac{7\times6}{2}(0.25)^5(0.75)^2 +$$
$$7(0.25)^6(0.75) + (0.25)^7$$

$$= 0.8665$$

$$P(X \geq 1) = 1 - P(X=0)$$
$$= 1 - {}^7C_0 (1/4)^0 (3/4)^7$$
$$= 1 - 0.13348$$
$$= 0.8665$$

Poisson's distribution:

Poisson's distribution can be used to approximate the binomial distribution.

The larger the $n$ and smaller the $p$, the better is the approximation.

If $X$ is the random variable denoting the number of occurrences in a given interval, for which the average rate of occurrences is $\lambda$ then, & according to the Poisson model, the probability of $x$ occurrences in that interval is given by.

$$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!} \qquad x = 0, 1, \ldots$$

where

$P(X=x)$ or $f(x;\lambda)$ = probability of $x$ successes given the parameters $n$ and $p$

$n$ = sample size

$p$ = probability of success

$e$ = mathematical constant approximated by 2.71828

$x$ = number of successes in the sample ($x = 0, 1, \ldots n$)
$\lambda = np$ is called as the rate, shape or intensity parameter.

Cummulative Poisson distribution:

A cummulative poisson probability refers to the probability that the Poisson random variable is greater than some specified lower limit and less than some specified upper limit.

The Poisson cummulative distribution function is given by:

$$F(x;\lambda) = \sum_{i=0}^{x} \frac{\lambda^i e^{-\lambda}}{i!}$$

Consider a telephone operator who, on the average handles five calls every 3 minutes. What is the probability that there will be
a) No calls in the next minute
b) At least two calls in the next minute?

Let $X$ = number of calls in a minute, then $X$ has a Poissons distribution with mean $\lambda = \frac{5}{3}$

a) $P(X=0) = \frac{\left(\frac{5}{3}\right)^0 e^{-5/3}}{0!} = e^{-5/3} = 0.189$

b) $P(X \geq 2) = 1 - P(X=0) - P(X=1)$

$= 1 - 0.189 - \frac{\left(5/3\right)^1 e^{-5/3}}{1!}$

$= 1 - 0.189 - 0.315$

$= 0.496$

Ten percent of the tool produced in a certain manufacturing process turn out to be defective. Find the probability that in a sample of 10 tools chosen at random exactly two will be defective by using
a) Binomial distribution b) poisson distribution

a) The probability of a defective tool is $p = 0.1$
Let $X$ be a random variable denoting the no. of defective tool out of 10 chosen.

$P(X=k) = {}^nC_k\, p^k q^{n-k}$

$P(X=2) = {}^{10}C_2\, p^2 q^{10-2}$

$p = 0.1$
$q = 0.9$

$= \frac{10 \times 9}{2} (0.1)^2 (0.9)^8$

$= 0.19371$

b) By poissons rule
$\lambda = np = 10 \times (0.1) = 1$

$P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$

$$P(X=2) = \frac{(1)^2 e^{-1}}{2!} = 0.18394$$

3. A manufacturer produces IC chips, 1 percent of which are defective. Find the probability that in a box containing 100 chips, no defectives are found.
(a) Using binomial approximation
(b) Using poisson approximation

(a) Binomial approximation:

$n = 100 \quad p = \frac{1}{100} = 0.01 \ (1\%) \quad x = 0$

$$b(x; n, p) = {}^nC_x \, p^x \, q^{n-x}$$

$$b(0; 100, 0.01) = {}^{100}C_0 \, (0.01)^0 \, (0.99)^{100}$$

$$= 1 \times 1 \times (0.99)^{100}$$

$$= 0.36603$$

Geometric distribution:

A random variable $X$ is said to follow geometric distribution if it assumes only non-negative values and its p.m.f is given by:

$$P(X=x) = pq^{x-1} \qquad x = 1, 2, \ldots$$

Mean and variance of geometric distribution:

If $X$ is a geometric random variable with parameter $p$, then the mean is given by

$$\mu = E(X) = \frac{1}{p}$$

and the variance of geometric distribution is

$$\sigma^2 = Var(X) = \frac{1-p}{p^2}$$

1. If the probability that a target is destroyed on any shot is 0.5. What is the probability that it would be destroyed on the 6th attempt?

$p = 0.5 \ ; \quad q = 1 - 0.5 = 0.5$

$$\therefore P(X=x) = pq^{x-1}$$

$$P(X=6) = (0.5)(0.5)^5 = 0.015625$$

Suppose that a trainee soldier shoots a target in an independent fashion. If the probability that the target is shot is 0.8,

(a) What is the probability that the target would be hit on the 6th attempt?

(b) What is the probability that it takes him less than 5 shots?

(c) What is the probability that it takes him an even number of shots?

Soln $p = 0.8 \quad q = 1 - p = 0.2$

(a) $x = 6$

Probability that the target would be hit on the 6th attempt.

$P(X=x) = pq^{x-1}$

$P(X=6) = (0.8)(0.2)^5 = 0.000256$

(b) Probability that it takes him less than 5 shots is

$P(X<5) = \sum_{k=1}^{4} pq^{k-1}$

$= (0.8)(0.2)^0 + (0.8)(0.2)^3 + (0.8)(0.2)^2 + (0.8)(0.2)^1$

$= 0.9984$

(c) The probability that it takes him an even number of shots

$= P(X=2) + P(X=4) + P(X=6) + \ldots$

$= (0.8)(0.2)^{2-1} + (0.8)(0.2)^{4-1} + (0.8)(0.2)^{6-1} + \ldots$

$= (0.8)(0.2)[1 + (0.2)^2 + (0.2)^4 + \ldots$

$= (0.2)(0.8)[1 + 0.04 + (0.04)^2 + \ldots$

$= (0.2)(0.8)[1 - 0.04]^{-1}$

$= (0.2)(0.8)(0.96)$

$= 0.1536$

3) An urn contains N white and M black balls. Balls are randomly selected one at a time until a black one is obtained. If we assume that each selected ball is replaced before the next one is drawn, what is the probability that

(a) Exactly n draws are needed and

(b) At least k draws are needed?

## Hypergeometric probability distribution:

The $x$ successes (defectives) can be choosen in $\left(C_x^m\right)$ ways.

The $n-x$ failure (non-defectives) can be choosen in $\left(C_{n-x}^{N-m}\right)$ ways

The $n$ objects can be choosen from a set of $N$ objects in $\left(C_n^N\right)$ ways.

If we consider all the posibilities as equally likely, it follows that for sampling without replacement the probability of getting $x$ successes in $n$ trials is

$$h(x; n, m, N) = \frac{C_x^m \; C_{n-x}^{N-m}}{C_n^N} \qquad x = 0, 1, \ldots n; \quad x \leq m; \; n-x \leq N-m$$

## Exponential distribution;

### Probability Density Function (pdf):

The probability density function (pdf) of an exponential distribution is

$$f_X(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The parameter $\lambda$ is called the rate parameter.

It is inverse of the expected duration ($\mu$).

$$\Rightarrow \lambda = \frac{1}{\mu}$$

### Cumulative distribution function:

The cumulative distribution function (cdf) of an exponential distribution is

$$F_X(x; \lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The CDF can also be written as the probability of the lifetime being less than some value $x$

$$P(X \leq x) = 1 - e^{-\lambda x}$$

### Mean and Variance of exponential distibution:

The expected value of an exponential random variable is $E[X] = \frac{1}{\lambda} = \mu$

The variance of an exponential random variable is

$$Var[X] = \frac{1}{\lambda^2} = \mu^2$$