# Probability

## Mutually Exclusive event:

If two (or) more of them can not happen simultaneously in the same trail then the events are called mutually exclusive events.

Eg: In throwing a dice experiment, the events 1, 2, 3, 4, 5, 6 are ME events.

## Equally likely Events

Outcomes of events are said to be equally likely.

Eg: Tossing a coin

## Probability

If a trial results in n-exhaustive, mutually exclusive and equally likely cases and m of them are favourable to happening of an event E then probability of an event E is denoted by P(E) and is defined as

$$\boxed{P(E) = \frac{\text{no. of favourable cases to event}}{\text{Total no. of Exhaustive events}} = \frac{m}{n}}$$

## AXIOMS OF PROBABILITY

1) $\boxed{0 \leq P(E) \leq 1}$ for event 'A' subset are equal to 'S'
2) $\boxed{P(S) = 1}$
3) If A and B are any two mutually exclusive events then $\boxed{P(A \cup B) = P(A) + P(B)}$

1) Find the probability of getting 1 red king if we select a card from a pack of 52 cards.

Sol: $n(A) = 2$

$n(S) = 52$

$P(E) = n(A)/n(S) = 2/52 = 1/26$

2) Three light bulbs are chosen at random from 12 bulbs of which 5 are defective. Find the probability that

i) All are defective

ii) one is defective

iii) Two are defective

Sol: No of Exhaustive cases $= {}^{12}C_3 = 220$

i) No of possible cases $= {}^5C_3 = 10$

$\therefore P(E) = 10/220 = \frac{1}{22}$

ii) No of possible cases $= {}^5C_1 \cdot {}^7C_2 = 105$

$\therefore P(E) = \frac{105}{220} = \frac{21}{44}$

iii) No of possible cases $= {}^5C_2 \cdot {}^7C_1 = 70$

$\therefore P(E) = \frac{70}{220} = \frac{7}{22}$

ELEMENTARY THEOREMS

1) $\boxed{\text{If } A, B \in F \text{ and } A \cap B = \phi,\ P(A \cup B) = P(A) + P(B)}$

We have, $A \cup B = A \cup B \cup \phi \ldots \cup \phi \ldots$

$\therefore P(A \cup B) = P(A) + P(B) + P(\phi) + \ldots + P(\phi)$ - (axiom 3)

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

2) $\boxed{P(A^c) = 1 - P(A)}$

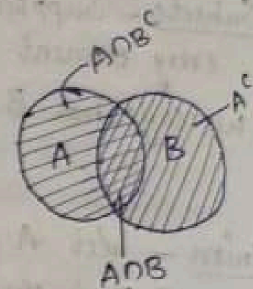
proof - W.K.T

$P(A \cup A^c) = P(S)$

$P(A) + P(A^c) = 1 \qquad \because A \cap A^c = \phi$

We have $\therefore P(A) = 1 - P(A^c)$

3) $\boxed{P(A \cap B^c) = P(A) - P(A \cap B)}$




proof - $(A \cap B^c) \cup (A \cap B) = A$

$\therefore P(A \cap B^c) \cup P(A \cap B) = P(A)$

$\Rightarrow P(A \cap B^c) + P(A \cap B) = P(A)$

$\Rightarrow P(A \cap B^c) = P(A) - P(A \cap B)$

Similarly

$P(A^c \cap B) = P(B) - P(A \cap B)$

4) $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$ $\Rightarrow$ Addition Theorem

proof - we have,

$A \cup B = (A^c \cap B) \cup (A \cap B) \cup (A \cap B^c)$

$\therefore P(A \cup B) = P(A^c \cap B) + P(A \cap B) + P(A \cap B^c)$

$= P(B) - \cancel{P(A \cap B)} + \cancel{P(A \cap B)} + P(A) - P(A \cap B)$

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

## Basics definitions

**Set** - A collection of well-defined objects is called a set.
The objects comprising the set are called elements.

$A = \{1, 2, 3, 5\}$

**Subsets** - Suppose A is a set, B is a set such that every element of B, belonging to the set A. We say that B is a subset of A and written as

$\boxed{B \subseteq A}$

**Union** - Let A and B be two sets. ~~The~~ Union of A and B is the set of all those elements which belonging to either A or B or both.

$A \cup B = \{x / x \in A \text{ or } x \in B\}$

**Intersection** - Let A and B be two sets. Intersection of A and B is the set of all those elements which are common to A and B

we write $A \cap B$

$A \cap B = \{x / x \in A \ \& \ x \in B\}$

**Null Set** - $\phi$ is the set which consists of no elements.

**Disjoint Set** - If $A \cap B = \phi$ then we say that A and B are disjoint.

**Universal Set** - All the sets are assumed to be subsets of some fixed set called the Universal set.

$\boxed{\therefore \phi \subset A \subset U}$

**Complement of A** :- The set of all elements which do not belong to A.

$A'$ or $A^c = \{x : x \in U / x \notin A\}$.

**Difference of sets** :- $A - B = \{x / x \in A \ \& \ x \notin B\}$

Note - $(A-B)$ & B are disjoint sets.

Eg - $A = \{1, 2, 3, 4, 5\}$, $B = \{2, 3, 6, 7, 8, 9, 10\}$

$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

$A \cap B = \{2, 3\}$

$A - B = \{1, 4, 5\}$

**De-Morgan's Law** :

$(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$

## Sample Space and Events

**Sample Space** - A set of all possible outcomes of an experiment is called a sample space

**Random Experiment**

If an experiment is conducted, any no. of times, under essentially identical conditions, there is a set of all possible outcomes associated with it. If the result is not certain and is anyone of the several possible outcomes, the experiment is called a random trial or a random experiment.

The outcomes are known as elementary events and a set of outcomes is an event.
(OR) Probabilistic situation is reffered to as a random experiment.

**Trial** – Each performance in a random experiment is called a trial i.e. all trials conducted under the same set of conditions from a random experiment.

**Outcome** – The result of a trial in a random experiment is called an outcome

**Event** – Every non-empty subset of a sample space of a random experiment is called an event

Eg: ① In tossing a coin is a trial, turning up Head or Tail is an outcome. There are two possible cases either Head or Tail.

$\therefore$ Sample Space = $\{H, T\}$

② In throwing a die, there are six possibilities

1, 2, 3, 4, 5, 6

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

Note – If a coin is tossed 'n' times or n coins tossed at a time the sample space consists of $2^n$ elements

**Exhaustive Events** – All possible events in any trial are known as Exhaustive events

**Permutation** – If r objects are chosen from a set of 'n' distinct objects, any particular arrangement or order of these objects is called a permutation. The no. of permutations of 'r' objects selected from a set of n objects is $\boxed{{}^nP_r = \frac{n!}{(n-r)!}}$

**Combination** – To find the no. of ways in which r objects can be selected from a set of n distinct objects is called the no. of combinations.

The no. of ways in which r objects can be selected from a set of n distinct objects is ${}^nC_r = \frac{n!}{(n-r)!\,r!}$

---

Eg 1: What is the probability that a card drawn at random from the pack of playing cards may be either a queen (or) King

Sol: $n(S) = {}^{52}C_1 = 52$

For Queen

$\therefore n(E_1) = {}^4C_1 = 4$

For King

$\therefore n(E_2) = {}^4C_1 = 4$

But $E_1, E_2$ are mutually exclusive events

$P(E_1 \cup E_2) = P(E_1) + P(E_2) = \frac{n(E_1)}{n(S)} + \frac{n(E_2)}{n(S)} = \frac{2}{13}$

Eg 2 – In a group there are 3 men and 2 women. Three persons are selected at random from this group. Find probability that one man and two women (or) two men and one women are selected.

Sol $n(S) = {}^5C_3 = 10$

$E_1$ be 1 man & 2 women $n(E_1) = {}^3C_1 \times {}^2C_2 = 3$

$E_2$ be 2 men & 1 woman $n(E_2) = {}^3C_2 \times {}^2C_1 = 6$

$E_1 \cup E_2$

$P(E_1 \cup E_2) = P(E_1) + P(E_2) = \frac{n(E_1)}{n(S)} + \frac{n(E_2)}{n(S)} = \frac{3}{10} + \frac{6}{10}$

$= \frac{9}{10}$

Eg3: A card is drawn from a well shuffled pack of cards. What is probability that it is either a spade or an ace.

Sol - $n(S) = 52$

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

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

Eg4- A problem is given to three persons P, Q, R whose respective chances of solving it are 2/7, 4/7, 4/9 respectively. What is probability that the problem is solved?

Sol - Probability of problem getting solved = 1 - (probability of none solved)

$= 1 - P(P \cup Q \cup R)^c$

$= 1 - P(P^c \cap Q^c \cap R^c)$

$= 1 - P(P^c)P(Q^c)P(R^c)$

$P(P) = \frac{2}{7} \Rightarrow P(P^c) = \frac{5}{7}$

$P(Q) = \frac{4}{7} \Rightarrow P(Q^c) = \frac{3}{7}$

$P(R) = \frac{4}{9} \Rightarrow P(R^c) = \frac{5}{9}$

$= 1 - \left(\frac{5}{7}\right)\left(\frac{3}{7}\right)\left(\frac{5}{9}\right)$

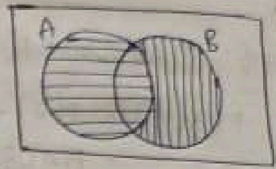
$= \cancel{1} \frac{122}{147}$

Theorem:

For any two events A and B : $\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$

Proof $A \rightarrow$ horizontal lines

$A^c \cap B \rightarrow$ Vertical lines




$A$ & $A^c \cap B$ are disjoint sets

$A$ and $A^c \cap B$ are mutually exclusive

$\therefore P[A \cup (A^c \cap B)] = P(A) + P(A^c \cap B)$

$A \cup B = A \cup (A^c \cap B)$

$P(A \cup B) = P(A \cup (A^c \cap B))$

$P(A \cup B) = P(A) + P(A^c \cap B)$

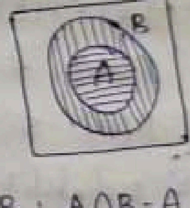
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Theorem:

If $A \subseteq B$ then prove that $\boxed{\text{i) } P(A^c \cap B) = P(B) - P(A) \quad \text{ii) } P(A) \leq P(B)}$

proof: i) $A \rightarrow$ horizontal lines

$B - A \rightarrow$ Vertical lines




$P(A^c \cap B) = P(B) - P(A \cap B)$

$= P(B) - P(A)$

Since $A \subseteq B$ ; $A \cap B = A$

ii) $P(B) - P(A) \geq 0$

$P(B) \geq P(A)$ ; $P(A) \leq P(B)$

Theorem: For any three events A, B and C

$$\boxed{P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)-P(C\cap A) + P(A\cap B\cap C)}$$

Proof: $P(A\cup B\cup C) = P(D\cup C)$ where $D = A\cup B$

$$P(D\cup C) = P(D) + P(C) - P(D\cap C)$$
$$= P(A\cup B) + P(C) - P((A\cup B)\cap C)$$
$$= P(A)+P(B)-P(A\cap B)+P(C)-P[(A\cap C)\cup(B\cap C)]$$

$$\boxed{P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C) + P(A\cap B\cap C)}$$

Theorem: If A and B are mutually exclusive events, then prove that $\boxed{P(A) \le P(B^c)}$

Sol: $P(A\cap B) = 0$

$$P(A\cup B) = P(A)+P(B)$$
$$P(A)+P(B) \le 1$$
$$P(A) \le 1-P(B)$$
$$P(A) \le P(B^c)$$

* Find the probability of getting one head in tossing two coins

Sol: $n(S) = 4 \quad n(A) = \{HT, TH\} = 2$

$$\therefore P(E) = \frac{n(A)}{n(S)} = \frac{2}{4} = \frac{1}{2}$$

* If $P(A\cup B) = \frac{4}{5}$, $P(B^c) = \frac{1}{3}$ and $P(A\cap B) = 1/5$

Find i, $P(B)$ ii, $P(A)$ iii, $P(A^c\cap B)$ iv, $P(A\cap B^c)$ v, $P(A^c\cap B^c)$ vi, $P(A^c\cup B^c)$

Sol: i, $P(B) = 1-P(B^c)$
$$= 1-\frac{1}{3}$$
$$P(B) = \frac{2}{3}$$

ii, $P(A\cup B) = P(A)+P(B)-P(A\cap B)$
$$\frac{4}{5} = P(A) + \frac{2}{3} - \frac{1}{5}$$
$$P(A) = 1-\frac{2}{3} = \frac{1}{3}$$

iii, $P(A^c\cap B) = P(B)-P(A\cap B)$
$$= \frac{2}{3}-\frac{1}{5}$$
$$= \frac{7}{15}$$

iv, $P(A\cap B^c) = P(A)-P(A\cap B)$
$$= \frac{1}{3}-\frac{1}{5} = \frac{2}{15}$$

v, $P(A^c\cap B^c) = P(A\cup B)^c$
$$= 1-P(A\cup B)$$
$$= 1-\frac{4}{5} = \frac{1}{5}$$

vi, $P(A^c\cup B^c) = P(A\cap B)^c$
$$= 1-P(A\cap B)$$
$$= 1-\frac{1}{5} = \frac{4}{5}$$

* If $B \subseteq A$ and A, B are two events such that $P(A) = 3P(B)$ and $A \cup B = S$ Find $P(B) = ?$

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

$P(A \cup B) = P(S) = 1$

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

$1 = 3P(B) + \cancel{P(B)} - \cancel{P(B)}$

$P(B) = 1/3$

* If A and B are two events $P(A) = 3/5$, $P(B) = 1/2$ then prove that

i, $P(A \cup B) \geq 3/5$ ii, $\frac{1}{10} \leq P(A \cap B) \leq 1/2$

Sol: i, $A \subseteq (A \cup B) \Rightarrow P(A) \leq P(A \cup B)$

$\frac{3}{5} \leq P(A \cup B) \Rightarrow P(A \cup B) \geq \frac{3}{5}$

ii, $A \cap B \leq B \Rightarrow P(A \cap B) \leq P(B)$

$P(A \cap B) \leq \frac{1}{2}$ — ①

$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq 1$

$\frac{3}{5} + \frac{1}{2} - P(A \cap B) \leq 1$

$\frac{3}{5} + \frac{1}{2} - 1 \leq P(A \cap B)$

$\frac{1}{10} \leq P(A \cap B)$ — ②

From ① & ②

$\frac{1}{10} \leq P(A \cap B) \leq \frac{1}{2}$

## Conditional Probability

Conditional Probability is calculating the probability of an event given that another event has already occured

$P(A/B)$ read as $P(A$ given $B)$ is

$P(A/B) = P(A \text{ and } B) / P(B)$

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

## Independent Events

Two events are said to be independent if the happening of an event is not affected in the happening of other event

→ $P(A/B)$ will be equal to $P(A)$ if they are independent events

→ $\boxed{P(A \cap B) = P(A) \cdot P(B)}$ where A and B are independent

Theorem: If A and B are independent events. The $A^c$ and $B^c$ are also independent events

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

$P(A^c \cap B^c) = P((A \cup B)^c)$

$= 1 - P(A \cup B)$

$= 1 - P(A) - P(B) + P(A \cap B)$

$= 1 - P(A) - P(B) + P(A)P(B)$

$= (1 - P(A)) - P(B)(1 - P(A))$

$= (1 - P(A))(1 - P(B))$

$= P(A^c)\, P(B^c)$

Theorem: If A & B are independent events. Then A and $B^c$ are also independent

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

$P(A\cap B^c) = P(A) - P(A\cap B)$

$= P(A) - P(A)P(B)$

$= P(A)(1 - P(B))$

$= P(A)P(B^c)$

Theorem: If A, B, C are mutually independent events. Then $A\cup B$ and C are also independent events

Proof: Given that A, B and C are independent events.

$\therefore P[(A\cup B)\cap C] = P[(A\cap C)\cup(B\cap C)]$

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

$= P(A)\cdot P(C) + P(B)\cdot P(C) - P(A)\cdot P(B)\cdot P(C)$

$= P(C)[P(A) + P(B) - P(A)\cdot P(B)]$

$= P(C)\,P(A\cup B)$

* In a class 40% of the students study math and science. 60% of the students study math. What is probability of a student studing science given he/she alread studying math?

Sol:

$P(M \text{ and } S) = 0.40$

$P(M) = 0.60$

$P(S/M) = \frac{P(M \text{ and } S)}{P(M)} = \frac{0.4}{0.6} = 0.67$

* There are 5 green, 7 red balls. Two balls are selected, one by one without replacment. Find the probability that first is green and second is red

Sol: $P(G)\times P(R) = \left(\frac{5}{12}\right)\left(\frac{7}{11}\right) = 35/132$

* Two aeroplanes bomb a target in succession. The probability of each correctly scaring a hit is 0.3 and 0.2 respectively. The second will bomb if the first misses the target. Find the probability that i, target is hit ii, both fails to score hit

Sol: Probability of 1st plane hitting target = $P(A) = 0.3$

Probability of 2nd plane hitting target = $P(B) = 0.2$

$P(\bar{A}) = 0.7 \quad P(\bar{B}) = 0.8$

i, P(target is hit) = P(A hits) or (A fails and B hits)

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

$= P(A) + P(\bar{A}\cap B)$ - (By addition Theorem)

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

$= 0.3 + (0.7)\times(0.2)$

$= 0.44$

ii, P(both fails) $= P(A^c\cap B^c)$

$= \cancel{P(A\cup B)}^c$

$= P(A^c)\cdot P(B^c)$

$= (0.7)\times(0.8)$

$= 0.56$

## Baye's Theorem

$E_1, E_2, E_3, \dots, E_n$ are n mutually exclusive and exhaustive events such that $P(E_i) > 0$ $(i = 1, 2, \dots n)$ in a sample space S and A is any other event in S intersecting with every $E_i$ i.e A can only occur in combination with any one of the events $E_1, E_2, \dots E_n$ such that $P(A) > 0$

If $E_k$ is any of the events of $E_1, E_2, \dots E_n$ where $P(E_1), P(E_2) \dots P(E_n)$ and $P(A|E_1), P(A|E_2) \dots P(A|E_n)$ are known then

$$\boxed{P(E_k/A) = \frac{P(E_k)\cdot P(A/E_k)}{P(E_1)\cdot P(A/E_1) + P(E_2)\cdot P(A/E_2) + \dots + P(E_n)\cdot P(A/E_n)}}$$

**Eg 1**: In a certain college, 25% of boys and 10% of girls are studying mathematics. The girls constitute 60% of the students

a) What is the probability that mathematics being studied

b) If a student is selected at random and is found to be studying mathematics, find the probability that the student is girl? c) a boy.

**Sol**: $P(Boy) = P(B) = \frac{40}{100} = \frac{2}{5}$

$P(Girls) = P(G) = \frac{60}{100} = \frac{8}{5}$

Probability that mathematics is studied by boy

$= P(M|B) = \frac{25}{100} = \frac{1}{4}$

Probability that mathematics is studied by girl

$= P(M/G) = \frac{10}{100} = \frac{1}{10}$

a) Probability that student studied maths $= P(M)$

$= P(G)\,P(M|G) + P(B)\,P(M|B)$

$= \frac{3}{5} \times \frac{1}{10} + \frac{2}{5} \times \frac{1}{4}$

$= \frac{3}{50} + \frac{2}{20} \Rightarrow \frac{4}{25}$

b) $P(G/M) = \frac{P(G)\,P(M/G)}{P(M)} = \frac{\frac{3}{5} \times \frac{1}{10}}{\frac{4}{25}} = \frac{3}{8}$

c) $P(B/M) = \frac{P(B)\cdot P(M/B)}{P(M)} = \frac{\frac{2}{5} \times \frac{1}{4}}{\frac{4}{25}} = \frac{10}{16} = \frac{5}{8}$

**Eg 2**. A Bag 'A' contains 2 white and 3 red balls and a Bag 'B' contains 4 white and 5 red balls : one ball is drawn at random from one of bags and it is found to be red. Find the probability that red ball drawn is from Bag B.

**Sol**: Let A and B denote the events of selecting Bag A and Bag B respectively.

$P(A) = \frac{1}{2}, P(B) = \frac{1}{2}$

Having selected Bag A, then probability to draw a red ball from A $= P(R/A) = \frac{3}{5}$

Similarly, $P(R/B) = \frac{5}{9}$

$$P(B/R) = \frac{P(B)\cdot P(R/B)}{P(A)\cdot P(R/A) + P(B)\cdot P(R/B)} = \frac{\frac{1}{2}\cdot\frac{5}{9}}{\frac{1}{2}\cdot\frac{3}{5} + \frac{1}{2}\cdot\frac{5}{9}}$$

$= \frac{25}{52}$

Eg 3: Of the three men, the chances that a politician, a business man or an academician will be appointed as a vice-chancellor (V.C) of a University are 0.5, 0.3, 0.2 respectively. Probability that research is promoted by these persons if they are appointed as V.C are 0.3, 0.7, 0.8 respectively.

i) Determine the probability that research is promoted

ii) If research is promoted, what is the probability that V.C is an academician?

Sol: Let A, B, C be the events that a politician, businessman or an academician will be appointed as V.C. of the three men.

$P(A) = 0.5$, $P(B) = 0.3$, $P(C) = 0.2$

The probabilities that research is promoted if they are appointed as V.C's are

$P(R/A) = 0.3$, $P(R/B) = 0.7$, $P(R/C) = 0.8$

i) The probability that the research is promoted

$P(R) = P(A) \cdot P(R/A) + P(B) \cdot P(R/B) + P(C) \cdot P(R/C)$

$= (0.5)(0.3) + (0.3)(0.7) + (0.2)(0.8)$

$= 0.52$

ii) The probability that research is promoted when the V.C is an academician.

$$P(C/R) = \frac{P(C) \cdot P(R/C)}{P(A) \cdot P(R/A) + P(B) \cdot P(R/B) + P(C) \cdot P(R/C)} \quad \frac{P(C) \cdot P(R/C)}{P(R)}$$

$$= \frac{(0.2)(0.8)}{(0.52)} = \frac{4}{13} = 0.30769$$

Eg 4: A businessman goes to hotels X, Y, Z, 20%, 50%, 30% of the time respectively. It is known that 5%, 4%, 8% of the rooms in X, Y, Z hotels have faulty plumbings. What is the probability that business man's room having faulty plumbing is assigned to hotel Z?

Sol: Let the probabilities of business man going to hotels X, Y, Z be respectively

$P(X) = \frac{20}{100} = \frac{2}{10}$, $P(Y) = \frac{50}{100} = \frac{5}{10}$, $P(Z) = \frac{30}{100} = \frac{3}{10}$

Let E be the event that the hotel room has faulty plumbing. Then the probabilities that hotels X, Y, Z have faulty plumbing are

$P(E/X) = \frac{5}{100} = \frac{1}{20}$, $P\left(\frac{E}{Y}\right) = \frac{4}{100} = \frac{1}{25}$, $P\left(\frac{E}{Z}\right) = \frac{8}{100} = \frac{2}{25}$

The probability that the business man's room having faulty plumbing is assigned to hotel Z

$$P\left(\frac{Z}{E}\right) = \frac{P(Z) \cdot P\left(\frac{E}{Z}\right)}{P(X) P\left(\frac{E}{X}\right) + P(Y) P\left(\frac{E}{Y}\right) + P(Z) P\left(\frac{E}{Z}\right)} = \frac{\frac{3}{10} \times \frac{2}{25}}{\frac{2}{10} \times \frac{1}{20} + \frac{5}{10} \times \frac{1}{25} + \frac{3}{10} \times \frac{2}{25}}$$

$$= \frac{4}{9}$$

Eg5: Suppose 5 men out of 100 and 25 women out of 10,000 are color blind. A color blind person is chosen at random. What is the probability of the person being male.

Sol: $P(M) = \frac{1}{2}$, $P(W) = \frac{1}{2}$

Let B represent a blind person then,

$P(B/M) = \frac{5}{100} = \frac{1}{20} = 0.05$, $P(B/W) = \frac{25}{10,000} = \frac{1}{400} = 0.0025$

The Probability that the choosen person is male i.e,

$$P\left(\frac{M}{B}\right) = \frac{P(M) \cdot P(B/M)}{P(M) \cdot P(B/M) + P(W) P(B/W)}$$

$$= \frac{0.05 \times 0.5}{(0.05 \times 0.5) + (0.5)(0.0025)} = 0.95$$

Eg6: In a Bolt factory machines A, B, C manufacture 20%, 30% and 50% of the total of their output and 6%, 3% and 2% are defective. A bolt is drawn at random and found to be defective. Find the probability that it is manufactured from i, Machine A ii, Machine B iii, Machine C.

Sol: $P(A) = \frac{20}{100} = \frac{1}{5}$, $P(B) = \frac{30}{100} = \frac{3}{10}$, $P(C) = \frac{50}{100} = \frac{1}{2}$

$P\left(\frac{D}{A}\right) = \frac{6}{100} = \frac{3}{50}$, $P(D/B) = \frac{3}{100}$, $P(D/C) = \frac{2}{100} = \frac{1}{50}$

If bolt is defective, then the probability that it is

i, from machine A

$$P(A/D) = \frac{P(A) \cdot P(D/A)}{P(A) \cdot P(D/A) + P(B) P(D/B) + P(C) P(D/C)}$$

$$P(A/D) = \frac{\frac{1}{5} \times \frac{3}{50}}{\frac{1}{5} \times \frac{3}{50} + \frac{3}{10} \times \frac{3}{100} + \frac{1}{2} \times \frac{1}{50}} = \frac{12}{31}$$

ii, from machine 'B'

$$P(B/D) = \frac{P(B) \cdot P(D/B)}{P(A) \cdot P(D/A) + P(B) \cdot P(D/B) + P(C) \cdot P(D/C)}$$

$$= \frac{\frac{3}{10} \times \frac{3}{100}}{\frac{1}{5} \times \frac{3}{50} + \frac{3}{10} \times \frac{3}{100} + \frac{1}{2} \times \frac{1}{50}} = \frac{9}{31}$$

iii, from machine 'C'

$$P(C/D) = \frac{P(C) \cdot P(D/C)}{P(A) \cdot P(D/A) + P(B) P(D/B) + P(C) \cdot P(D/C)}$$

$$= \frac{\frac{1}{2} \times \frac{1}{50}}{\frac{1}{5} \times \frac{3}{50} + \frac{3}{10} \times \frac{3}{100} + \frac{1}{2} \times \frac{1}{50}} = \frac{10}{31}$$

Eg7: Companies $B_1$, $B_2$, $B_3$ produce 30%, 45% and 25% of the cars respectively. It is known that 2%, 3% and 2% of the cars produced from $B_1$, $B_2$, $B_3$ are defective. If a car purchased is found to be defective What is the Probability that this car is produced by $B_3$

Sol: $P(B_1) = \frac{30}{100} = \frac{3}{10}$, $P(B_2) = \frac{45}{100} = \frac{9}{20}$, $P(B_3) = \frac{25}{100} = \frac{1}{4}$

$P(D/B_1) = \frac{2}{100}$, $P(D/B_2) = \frac{3}{100}$, $P(D/B_3) = \frac{2}{100}$

If car purchased is found to be defective. The probability that the car is produced by company $B_3$?

$$P\left(B_3/D\right) = \frac{P(B_3)\,P(D/B_3)}{P(B_1)\,P(D/B_1) + P(B_2)P(D/B_2) + P(B_3)P(D/B_3)}$$

$$= \frac{\frac{1}{4} \times \frac{2}{100}}{\frac{3}{10} \times \frac{2}{100} + \frac{9}{20} \times \frac{3}{100} + \frac{1}{4} \times \frac{2}{100}}$$

$$= \frac{10}{49}$$

## Discrete Probability - Unit 2

1. **Discrete Random Variable**: A random variable x which can take only a finite number of discrete values in an interval of domain is called a discrete random variable. In other words, if the random variable takes the values only on the set $\{0,1,2 \ldots n\}$ is called a discrete random variable.

Eg: Tossing a coin, throwing a dice, the number of defective in a sample of electric bulbs, the no. of printing mistakes in each page of a book, the no. of telephone calls received by the telephone ~~~~ operator are examples of Discrete Random variable.

A few examples are

i. In the example (1), $X(S) = \{s : s = 0,1,2\}$ or Range of $X = \{0,1,2\}$

The random variable x is a discrete random variable

ii. The random variable denoting the no. of students in a class is

$X(x) = \{x : x \text{ is a positive integer}\}$

### Introduction

Suppose S is the sample space of some experiment. We know that outcomes of the experiments the elements of the sample space S and they need not be numbers. Sometimes we wish to assign a specific number to each outcome.

Eg: No. of heads in tossing 2 or 3 coins

Sum of the points on a pair of dice when they are thrown.

### Random Variables:

A random variable X on a sample space S is a function form S into the set R of real numbers such that the pre image of every element of it is an event of S

### Types of Random Variables

Random Variable is of two types

i. Discrete Random Variable

ii. Continous Random Variable

then the function $P(x)$ is called the probability mass function of the random variable X and the set

$\{P(x_i)\}$, $i = 1, 2, 3 \dots$ is called the discrete probability distribution of the discrete ~~var~~ random variable $x$.

The probability distribution of the random variable $x$ is given by mean of the following table

| $x$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $\dots$ | $x_i$ | $\dots$ | $x_n$ |
|---|---|---|---|---|---|---|---|---|
| $P(x)$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $\dots$ | $P_i$ | $\dots$ | $P_n$ |

Further $P(x < x_i) = P(x_1) + P(x_2) + \dots + P(x_{i-1})$

$P(x \le x_i) = P(x_1) + P(x_2) + \dots + P(x_{i-1}) + P(x_i)$

and $P(x > x_i) = 1 - P(x \le x_i)$

## Continous Random Variable:

A Random Variable $x$ which can take values continuously i.e, which takes all possible values in a given interval is called a continuous random variable.

For example, the height, age and weight of individuals are examples of continuous random variable. Also temperature and time are continous random variables.

## Probability Function of a Discrete Random Variable

For a discrete random variable $x$, the real valued function $P(x)$ is such that $P(X = x) = P(x)$ the $P(x)$ is called probability function or probability density function of a discrete random variable $x$. probability function $p(x)$ gives the measure of probability for different values of $x$.

## Properties of a probability function

If $P(x)$ is a probability function of a random variable $x$, then it possesses the following properties:

Similarly, $P(5) = P(x=5)$

$$= P\begin{pmatrix}(1,5),(5,1),(2,5),(5,2),(3,5),(5,3),\\(4,5),(5,4),(5,5)\end{pmatrix}$$

$$= \frac{9}{36}$$

and $P(6) = P(x=6)$

$$= P\begin{pmatrix}(1,6),(6,1),(2,6),(6,2),(3,6),(6,3),(4,6),\\(6,4),(5,6),(6,5),(6,6)\end{pmatrix}$$

$$= \frac{11}{36}$$

$\therefore$ The required discrete probability distribution is

| $X = x_i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(X = x_i) = P(x_i)$ | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

i, Mean, $\mu = \sum_{i=1}^{6} P_i x_i = 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right)$

$= 4.47$

i.e, $10k^2 + 9k - 1$ i.e, $(10k-1)(k+1) = 0$

$\therefore k = \frac{1}{10} = 0.1$ (Since $P(x) \ge 0$, So $k \ne -1$)

ii, $P(x \le 6) =$ $P(x=0) + P(x=1) + \dots + P(x=5)$

$= 0 + k + 2k + 2k + 3k + k^2 = 8k + k^2$

$= 0.81$ ($\because k = 0.1$)

$P(x \geq 6) = 1 - P(x < 6) = 1 - 0.81 = 0.19$

$P(0 < x < 5) = P(x=1) + P(x=2) + P(x=3) + P(x=4)$

$= k + 2k + 2k + 3k = 8k = \frac{8}{10} = 0.8$

$P(0 \leq x \leq 4) = P(x=0) + P(x=1) + P(x=2) + P(x=3) + P(x=4)$

$= 0 + k + 2k + 2k + 3k = 8k = 0.8$

iii, The required minimum value of $k$ is obtained as below

$P(x < 1) = P(x=0) + P(x=1) = 0 + k = 0.1$

$P(x \leq 2) = [P(x=0) + P(x=1)] + P(x=2)$

---

**Eg1:** Two dice are thrown. Let $x$ assign to each point $(a,b)$ in $S$ the minimum of its numbers i.e, $x(a,b) = \max(a,b)$. Find the probability distribution $x$ is a random variable with $x(S) = \{1,2,3,4,5,6\}$. Also find the mean and variance of the distribution

(or) A random variable $x$ has the following distribution?

| $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(x)$ | 1/36 | 3/36 | 5/36 | 7/36 | 9/36 | 11/36 |

Find (a) mean (b) Variance (c) $P(1 < x < 6)$

**Solution**

The total number of cases are $6 \times 6 = 36$

The maximum number could be $1,2,3,4,5,6$ i.e $x(S) = x(a,b) = \max(a,b)$

The number 1 will appear only in one case $(1,1)$ So $p(1) = p(x=1) = p(1,1) = 1/36$

For maximum 2, favourable cases are $(1,2), (2,1), (2,2)$

So $P(2) = P(x=2) = 3/36$

For maximum 3, favourable cases are $(1,3),(3,1),(2,3),(3,2),(3,3)$ So $P(3) = P(x=3) = 5/36$

Illy $P(4) = P(x=4) = 7/36$

$P(5) = P(x=5) = 9/36$

$P(6) = P(x=6) = 11/36$

(a) Mean :

$\mu = 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right)$

$= \frac{1}{36}(1 + 6 + 15 + 28 + 45 + 66)$

$= 4.47$

(b) Variance :

$\sigma = \sum_{i=1}^{6} P_i(x_i^2) - (\mu)^2$

$= \left[1^2\left(\frac{1}{36}\right) + 2^2\left(\frac{3}{36}\right) + 3^2\left(\frac{5}{36}\right) + 4^2\left(\frac{7}{36}\right) + 5^2\left(\frac{9}{36}\right) + 6^2\left(\frac{11}{36}\right)\right] - (4.47)^2$

$= \left[\frac{1}{36}(1 + 12 + 45 + 112 + 225 + 396)\right] - (4.47)^2$

$= \frac{791}{36} - (4.47)^2 = 1.99$

(c) $P(1 < x < 6) = P(X=2) + P(x=3) + P(x=4) + P(x=5)$

$= \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36}$

$= \frac{24}{36} = \frac{4}{6} = \frac{2}{3}$

~~Eg 2 A random variable x has the following probability function.~~

Eg 2: Let $x$ denotes the sum of the two numbers that appear when a pair of fair dice is tossed. Determine the i, distribution function ii, mean and iii, variance (or) Two dice are rolled at random. Find the probability distribution of the sum of the numbers on them. Also find the mean of the distribution

Sol: If two unbiased dice are thrown, then the sum $x$ of the two numbers which turn up must be an integer between 2 and 12

For $x=2$, there is only one favourable point $(1,1)$

So, $P(X=2) = 1/36$.

For $x=3$, there are two favourable points $(1,2)\ (2,1)$

So, $P(X=3) = 2/36$

Illy
$P(X=4) = 3/36$ $\quad P(X=7) = 6/36$ $\quad P(X=10) = 3/36$

$P(X=5) = 4/36$ $\quad P(X=8) = 5/36$ $\quad P(X=11) = 2/36$

$P(X=6) = 5/36$ $\quad P(X=9) = 4/36$ $\quad P(X=12) = 1/36$

ii, Mean,

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $P(x)$ | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

Mean: $\mu = 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + 4\left(\frac{3}{36}\right) + \cdots + 12\left(\frac{1}{36}\right)$

$\mu = \frac{252}{36} = 7$

* The probability density function of a variable $x$ is

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $P(x)$ | K | 3K | 5K | 7K | 9K | 11K | 13K |

i, Find K ii, $P(x<4)$; $P(x \geq 5)$.
iii, What will be the minimum value of K so, that $P(x \leq 2) > 0.3$

Sol: i, $\sum_{x=0}^{6} P(x_i) = 1$

$K + 3K + 5K + 7K + 9K + 11K + 13K = 1$

$K = \frac{1}{49}$

ii, $P(x<4) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$= K + 3K + 5K + 7K = 16K$

$= 16/49$

$P(x \geq 5) = P(x=5) + P(x=6)$

$= 11K + 13K = 24K = \frac{24}{49}$

iii, $P(x \leq 2) = P(x=0) + P(x=1) + P(x=2) > 0.3$

$K + 3K + 5K > 0.3$

$9K > 0.3$

$K > \frac{0.3}{9}$

$K > \frac{3}{90} \Rightarrow K > \frac{1}{30}$

$\therefore$ minimum value of $K = 1/30$

* A random variable $x$ has the following probabilities functions

| $x_i$ | $-3$ | $-2$ | $-1$ | $0$ | $1$ | $2$ | $3$ |
|---|---|---|---|---|---|---|---|
| $P(x_i)$ | $k$ | $0.1$ | $k$ | $0.2$ | $2k$ | $0.4$ | $2k$ |

Find i, k (ii) mean (iii) variance

Sol: i) $\sum P(x_i) = 1$

$$0.2 + k + 0.1 + k + 2k + 0.4 + 2k = 1$$

$$\cancel{4k + 0.6 = 1} \qquad 6k + 0.7 = 1$$

$$\cancel{4k = 0.4} \qquad \boxed{k = 1/20}$$

$$\cancel{k = 0.1}$$

ii, Mean $= \sum_{i=0}^{n} p_i x_i$

$$\mu = -3(k) + (-2)(0.1) + (-1)(k) + 0(0.2) + 1(2k) + 2(0.4) + 3(2k)$$

$$= 4k + 0.6 = \frac{4}{20} + 0.6 = 0.2 + 0.6 = 0.8$$

$$\boxed{\mu = 0.8}$$

iii, Variance

$$\sigma^2 = E(x^2) - [E(x)]^2 = \sum_{i=0}^{n} p_i x_i^2 - \mu^2$$

$$\sigma^2 = (9)(k) + 4(0.1) + 1(k) + 0(0.2) + 1(2k) + 4(0.4) + 9(2k) - (0.8)^2$$

$$\sigma^2 = (30k + 2) - 0.64$$

$$\boxed{\sigma^2 = 2.86}$$

## Continous Probability distribution

When a random variable $x$ takes every value in an interval, it give rise to continous distribution of $\bar{x}$. Consider the small interval $\left[x - \frac{dx}{2}, x + \frac{dx}{2}\right]$ of length $dx$ round the point $x$. Let $f(x)$ be any continuous function of $x$ so that $f(x)\,dx$ represents the probability that the variable $x$ falls in the infinitesimal interval $\left[+x - \frac{dx}{2}, x + \frac{dx}{2}\right]$

It can be represented as

$$P\left(x - \frac{dx}{2} \le x \le x + \frac{dx}{2}\right) = f(x)\,dx.$$ Then $f(x)$ is called the probability density function

Properties of probability density function $f(x)$:

i) $f(x) \ge 0 \;\forall\, x \in R$ ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

iii) The probability $P(E)$ is given by

$P(E) = \int_E f(x)\,dx$ is well defined for event $E$

**Mean:**

$$\mu = E(x) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$$

**Variance:**

$$\sigma^2 = \int_{-\infty}^{\infty} E(x^2)\,dx - [E(x)]^2$$

(or)

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

**Eg 1:** If the probability density of a random variable is given by $f(x) = \begin{cases} k(1-x^2) & ; \text{for } 0 < x < 1 \\ 0 & ; \text{otherwise} \end{cases}$

i, find the value of $k$ ii, b/w 0·1 and 0·2

**Sol:** W.K.T $\int_{-\infty}^{\infty} f(x)\,dx = 1$

i.e $\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx = 1$

$0 + \int_{0}^{1} f(x)\,dx + 0 = 1$

$\int_{0}^{1} k(1-x^2)\,dx = 1$

$\Rightarrow k \int_{0}^{1} (1-x^2)\,dx = 1 \Rightarrow k\left[x - \frac{x^3}{3}\right]_0^1 = 1$

$\Rightarrow k\left[\left(1-\frac{1}{3}\right) - (0-0)\right] = 1,$

$k - \frac{k}{3} = 1 \Rightarrow 2k = 3 \Rightarrow \boxed{k = 3/2}$

ii, $\int_{0.1}^{0.2} k(1-x^2)\,dx = \frac{3}{2}\int_{0.1}^{0.2} (1-x^2)\,dx$

$= \frac{3}{2}\left[x - \frac{x^3}{3}\right]_{0.1}^{0.2} = \frac{3}{2}\left[(0.2-0.1) - \left(\frac{(0.2)^3}{3} - \frac{(0.1)^3}{3}\right)\right]$

$= 0.2965$

**Eg 2:** If a random variable has the probability density $f(x)$ as $f(x) = \begin{cases} 2e^{-2x} & ; \text{for } x > 0 \\ 0 & ; \text{for } x \le 0 \end{cases}$

Find the probability that it will take on a value.
i, b/w 1 and 3 ii, greater than 0·5

**Sol:** i, The probability that a variate takes a value b/w 1 and 3 is given by

$P(1 \le x \le 3) = \int_1^3 f(x)\,dx = \int_1^3 2e^{-2x} \Rightarrow 2\left[\frac{e^{-2x}}{-2}\right]_1^3$

$= -\left[e^{-6} - e^{-2}\right] = e^{-2} - e^{-6}$

ii, The probability that a variable takes a value greater than 0·5 is

$P(x \ge 0.5) = \int_{0.5}^{\infty} f(x)\,dx = \int_{0.5}^{\infty} 2e^{-2x}\,dx$

$\Rightarrow 2\left[\frac{e^{-2x}}{-2}\right]_{0.5}^{\infty} = -\left[e^{-2(\infty)} - e^{-2(0.5)}\right] = -\left[0 - e^{-1}\right]$

$\Rightarrow e^{-1} = 1/e$

**Eg 3:** A continous random variable $x$ has the distribution function $f(x) = \begin{cases} 0 & ; \text{if } x \le 1 \\ 4k(x-1)^3 & ; \text{if } 1 \le x \le 3 \\ 0 & ; \text{if } x > 3 \end{cases}$

Determine i, $k$ ii, Mean.

**Sol:** i, Since total probability is Unity.

$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{1} f(x)\,dx + \int_{1}^{3} f(x)\,dx + \int_{3}^{\infty} f(x)\,dx = 1$

$= 0 + \int_1^3 4k(x-1)^3\,dx + 0 = 1$

$4k\left[\frac{(x-1)^4}{4}\right]_1^3 = 1 \Rightarrow 4k\left[\frac{16-0}{4}\right] = 1$

$\Rightarrow 16k = 1 \Rightarrow \boxed{k = 1/16}$

ii, Mean :

$$\mu = \int_1^3 x f(x)\,dx = \int_1^3 x\left(4k\,(x-1)^3\right)dx$$

$$= 4\times\frac{1}{16_4}\int_1^3 x(x-1)^3\,dx$$

$$= \frac{1}{4}\int_1^3 x\left[x^3+1-3x^2+3x\right]dx$$

$$= \frac{1}{4}\int_1^3 \left(x^4+x-3x^3+3x^2\right)dx$$

$$= \frac{1}{4}\left[\left(\frac{x^5}{5}\right)_1^3 - 3\left(\frac{x^4}{4}\right)_1^3 + 3\left(\frac{x^3}{3}\right)_1^3 + \left(\frac{x^2}{2}\right)_1^3\right]$$

$$= \frac{1}{4}\left[\frac{3^5-1}{5} - 3\left(\frac{3^4-1}{4}\right) + 3\left(\frac{3^3-1}{3}\right) - \frac{1}{2}\left(3^2-1\right)\right]$$

$$= 2.6$$

---

Eg4: A continuous random variable has the probability density function $f(x) = \begin{cases} kxe^{-\lambda x} & : \text{for } x\geq 0,\ \lambda>0 \\ 0 & : \text{otherwise} \end{cases}$

Determine i, k ii, Mean iii, Variance.

Sol : i, Since the total probability is unity, we have

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 = \int_{-\infty}^{0} f(x)\,dx + \int_0^{\infty} f(x)\,dx = 1$$

$$= 0 + \int_0^{\infty} f(x)\,dx = 1 \Rightarrow \int_0^{\infty} kxe^{-\lambda x}\,dx = 1$$

$$= k\left[\frac{x\,e^{-\lambda x}}{-\lambda} - 1\cdot\frac{e^{-\lambda x}}{\lambda^2}\right]_0^{\infty} = 1 \quad \left[\int u\cdot v\,dx = uv_1 - u'v_{11} + \dots\right]$$

$$= k\left[(0-0) - (0 - 1/\lambda^2)\right] = 1$$

$$\Rightarrow k\left(1/\lambda^2\right) = 1 \Rightarrow \boxed{k=\lambda^2}$$

ii, Mean :

$$\mu = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{0} x f(x)\,dx + \int_0^{\infty} x f(x)\,dx$$

$$\mu = \int_0^{\infty} x\left[kxe^{-\lambda x}\right]dx = \lambda^2\int_0^{\infty} x^2 e^{-\lambda x}\,dx$$

$$\mu = \lambda^2\left[x^2\frac{e^{-\lambda x}}{-\lambda} - 2x\frac{e^{-\lambda x}}{\lambda^2} + 2\frac{e^{-\lambda x}}{-\lambda^3}\right]_0^{\infty}$$

$$\mu = \lambda^2\left[0 - \frac{2}{-\lambda^3}\right] = \lambda^2\left[\frac{2}{\lambda^3}\right]$$

$$\boxed{\mu = \frac{2}{\lambda}}$$

iii, Variance : $\boxed{\int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2}$

$$\sigma^2 = \left[\int_{-\infty}^{0} x^2 f(x)\,dx + \int_0^{\infty} x^2 f(x)\,dx\right] - \mu^2$$

$$\sigma^2 = 0 + \int_0^{\infty} x^2\left[kxe^{-\lambda x}\right]dx - \mu^2$$

$$\sigma^2 = \lambda^2\int_0^{\infty} x^3 e^{-\lambda x}\,dx - \mu^2$$

$$\sigma^2 = \lambda^2\left[x^3\frac{e^{-\lambda x}}{-\lambda} - 3x^2\frac{e^{-\lambda x}}{\lambda^2} + 6x\frac{e^{-\lambda x}}{-\lambda^3} - \frac{6e^{-\lambda x}}{\lambda^4}\right]_0^{\infty} - \mu^2$$

$$\sigma^2 = \lambda^2\left[0 - \left(-\frac{6}{\lambda^4}\right)\right] - \frac{4}{\lambda^2} = \lambda^2\left(\frac{6}{\lambda^4}\right) - \frac{4}{\lambda^2}$$

$$\boxed{\sigma^2 = \frac{2}{\lambda^2}}$$

**Eg5:** For the continuous random variable x whose probability density function $f(x) = \begin{cases} c\,x(2-x) ; & \text{if } 0 \le x \le 2 \\ 0 ; & \text{if otherwise} \end{cases}$

wher c is constant find i, c ii, mean iii, variance

**Sol** i, $\int_{-\infty}^{\infty} f(x)\,dx = 1$ [Since Total Probability is unity]

$\Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx + \int_{2}^{\infty} f(x)\,dx = 1$

$0 + \int_0^2 c(2x - x^2)\,dx + 0 = 1$

$\Rightarrow c\left[2\left(\frac{x^2}{2}\right)_0^2 - \left(\frac{x^3}{3}\right)_0^2\right] = 1 \Rightarrow c\left[2(2) - \frac{8}{3}\right] = 1$

$\Rightarrow c\left[\frac{12-8}{3}\right] = 1 \Rightarrow c\left(\frac{4}{3}\right) = 1 \Rightarrow \boxed{c = 3/4}$

ii, Mean: $\mu = \int_0^2 x f(x)\,dx$

$\mu = \int_0^2 x\left[c\,x(2-x)\,dx\right] \Rightarrow \frac{3}{4}\int_0^2 (2x^2 - x^3)\,dx = \frac{3}{4}\left[2\left(\frac{x^3}{3}\right)_0^2 - \left(\frac{x^4}{4}\right)_0^2\right]$

$\mu = \frac{3}{4}\left[\frac{16}{3} - \frac{16}{4}\right] = \frac{3}{4}\left[16\left(\frac{1}{12}\right)\right] = 1 \quad \boxed{\mu = 1}$

iii, Variance $= \int_0^2 x^2 f(x)\,dx - \mu^2$

$\sigma^2 = \int_0^2 x^2\left[c\,x(2-x)\right]dx - \mu^2 = \frac{3}{4}\int_0^2 (2x^3 - x^4)\,dx - \mu^2 = \frac{3}{4}\left[2\left(\frac{x^4}{4}\right)_0^2 - \left(\frac{x^5}{5}\right)_0^2\right] - \mu^2$

$\sigma^2 = \frac{3}{4}\left[\frac{32}{4} - \frac{32}{5}\right] - 1^2 = \frac{3}{4}\left[32\left(\frac{1}{20}\right)\right] - 1 = \frac{24}{20} - 1$

$\sigma^2 = \frac{6-5}{5} = 1/5$

$\boxed{\sigma^2 = 1/5}$

**Eg6:** Probability density function of a random variable x is $f(x) = \begin{cases} \frac{1}{2}\sin x ; & \text{for } 0 \le x \le \pi \\ 0 ; & \text{otherwise} \end{cases}$ Find mean, variance.

**Sol.** i, Mean: $\mu = \int_{-\infty}^{\infty} f(x)\,dx$

$\mu = \int_{-\infty}^{0} x f(x)\,dx + \int_0^{\pi} x f(x)\,dx + \int_{\pi}^{\infty} x f(x)\,dx$

$\mu = 0 + \int_0^{\pi} x\left[\frac{1}{2}\sin x\right]dx + 0$

$\mu = \frac{1}{2}\int_0^{\pi} x\sin x\,dx = \frac{1}{2}\left[x(-\cos x) - (-\sin x)\right]_0^{\pi}$

$\mu = \frac{1}{2}\left[\left[\pi(-\cos\pi + \cos 0) + (\sin\pi - \sin 0)\right] - (0-0)\right]$

$\mu = \frac{1}{2}\left[\pi(1+1)\right] = \mu = \frac{2\pi}{2} \Rightarrow \boxed{\mu = \pi}$

ii, Variance: $\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$

$\sigma^2 = \int_{-\infty}^{0} x^2 f(x)\,dx + \int_0^{\pi} x^2 f(x)\,dx + \int_{\pi}^{\infty} x^2 f(x)\,dx - \mu^2$

$\sigma^2 = \int_0^{\pi} x^2\left[\frac{1}{2}\sin x\right]dx - \pi^2$

$\sigma^2 = \frac{1}{2}\left[x^2(-\cos x) - 2x(-\sin x) + 2(\cos x)\right]_0^{\pi} - \pi^2$

$\sigma^2 = \frac{1}{2}\left[\pi^2(1+1) - 0 + 2(-1-1)\right] - \pi^2$

$\sigma^2 = \frac{1}{2}\left[2\pi^2 - 4\right] - \pi^2$

$\sigma^2 = -2$

# BINOMIAL DISTRIBUTION

Binomial distribution was discovered by James Bernoli in the year 1700 and it is discrete probability distribution.

A random variable x has a binomial distribution if it assumes only non-negative values and its probability distribution is given by

$$P(x=r) = P(r) = \begin{cases} {}^nC_r\, p^r q^{n-r} & ; r = 0,1,2,\dots,n; q = 1-p \\ 0 & ; \text{otherwise} \end{cases}$$

1) A fair coin is tossed six times. Find the probability of getting four heads

Sol:

p = probability of getting a head = 1/2

q = probability of not getting head = 1/2

and $n = 6$, $r = 4$

$$\therefore P(x=r) = P(x=4) = P(4) = {}^6C_4 \left(\tfrac{1}{2}\right)^4 \left(\tfrac{1}{2}\right)^{6-4}$$

$$= \frac{6!}{4!(6-4)!}\left(\tfrac{1}{2}\right)^6$$

$$= \frac{6!}{4!\,2!}\left(\tfrac{1}{2}\right)^6$$

$$= \frac{\overset{3}{\cancel{6}}\times 5\times \cancel{4!}}{\cancel{4!}\times \cancel{2}}\left(\tfrac{1}{2}\right)^6$$

$$= \frac{15}{64}$$

2) A die is thrown 6 times. If getting an even number is a success, find the probability of

i, at least one success. ii, ≤ 3 success iii, 4 success

Sol

$P(e) = 1/2$, $q(e) = 1/2$

i, $P(r \geq 1) = 1 - P(r=0)$

$$= 1 - \left({}^6C_0 (1/2)^{6-0} (1/2)^0\right)$$

$$= 1 - (1/2)^6 = 1 - 1/64 = \frac{63}{64}$$

ii, $P(r \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$

$$= (1/2)^6 \left[{}^6C_0 + {}^6C_1 + {}^6C_2 + {}^6C_3\right] = \frac{21}{32}$$

iii, $P(r=4) = {}^6C_4 \left(\tfrac{1}{2}\right)^4 \left(\tfrac{1}{2}\right)^{6-4}$

$$= 15 \cdot 1/2^4 \cdot 1/2^2 = \frac{15}{2^6} = \frac{15}{64}$$

3) If 3 of 20 tyres are defective and 4 of them are randomly choosen for inspection, what is the probability that only one of the defective tyre will be included?

Sol

p = probability of defective tyre (success) = 3/20

$q = 1 - p = \frac{17}{20}$

Given $n = 4$

$$P(r=1) = {}^4C_1 \left(\frac{3}{20}\right)^1 \left(\frac{17}{20}\right)^{4-1}$$

$$= \frac{4\times 3\times (17)^3}{(20)^4} = 0.3685$$

**4)** Determine the binomial distribution for which the mean is 4 and variance 3.

**Sol.** Given mean $= 4$ i.e $np = 4$

variance $= 3$ i.e $npq = 3$

$(4)q = 3 \Rightarrow q = 3/4 \Rightarrow p = 1-q$

$p = 1/4$

$np = 4 \Rightarrow n \times 1/4 = 4$

$\Rightarrow n = 16$

---

**5)** The mean and variance of a binomial distribution are 4 and 4/3 respectively. Find $P(x \geq 1)$

**Sol:** $np = 4$ , $npq = 4/3$

$(4)q = 4/3$

$q = 1/3 \Rightarrow p = 1 - q = 1 - \frac{1}{3} = \frac{2}{3}$

$n \times \frac{2}{3} = 4^2$

$n = 6$

$P(x \geq 1) = 1 - P(x = 0)$

$= 1 - {}^6C_0\, p^0 q^{6-0}$

$= 1 - (1/3)^6 = 0.99$

---

**6)** In eight throws of a die, 5 (or) 6 is considered as success. Find its mean and variance & s.d

**Sol:** $p$ = probability of success $= 1/6 + 1/6 = \frac{2}{6} = 1/3$

$q$ = probability of failure $= 1 - p = \frac{2}{3}$

$n$ = no. of throws $= 8$

mean $= np = \frac{8}{3}$ , variance $= npq = 8\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)$

$= \frac{16}{9}$

$S.d = \sqrt{\text{variance}} = \sqrt{\frac{16}{9}} = \frac{4}{3}$

---

**7)** Out of 800 families with 5 children each, how many you expect to have a) 3 boys b) 5 girls c) either 2 (or) 3 boys d) atleast one boy ? Assume equal probabilities for boys and girls

**Sol:** Let no. of boys in each family $= x$

$p$ = probability of each boy $= 1/2$ , $q = 1 - 1/2 = 1/2$.

no. of children : $n = 5$

The probability distribution is

$P(x = r) = {}^nC_r\, p^r q^{n-r} = {}^5C_r (1/2)^r \cdot (1/2)^{5-r} = {}^5C_r (1/2)^5$

a) $P(x = 3) = {}^5C_3 \left(\frac{1}{2}\right)^5 = \frac{10}{32} = 5/16$ per family

Thus for 800 families the probability of no. of families having 3 boys $= \frac{5}{16} \times 800 = 250$ families

b) $P(x = 0) = P(5\text{ girls}) = \frac{1}{2^5}\, {}^5C_0 = 1/32$ per family

Thus for 800 families $= 1/32 \times 800 = 25$ families

c) $P(\text{either } 2 \text{ (or) } 3 \text{ boys}) = P(x=2) + P(x=3)$

$= \frac{1}{2^5}\left({}^5C_2 + {}^5C_3\right) = \frac{20}{32} = \frac{5}{8}$

Thus for 800 families $= \frac{5}{8} \times 800 = 500$ families

d) $P(\text{Atleast one boy}) = P(x=1) + P(x=2) + P(x=3) + P(x=4) + P(x=5)$

$= 1/2^5 \left({}^5C_1 + {}^5C_2 + {}^5C_3 + {}^5C_4 + {}^5C_5\right)$

$= 1/2^5 (5 + 10 + 10 + 5 + 1) = \frac{31}{32}$

Thus for 800 families $= \frac{31}{32} \times 800 = 775$ families

# POISSON DISTRIBUTION

A random variable $x$ is said to follow a poisson distribution if it assumes only non-negative values and its probability distribution is given by

$$P(x,\lambda) = P(X=x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & ; x = 0,1,2\ldots \\ 0 & ; \text{otherwise} \end{cases}$$

Here $\lambda > 0$ is called the parameter of the distribution.

1) A hospital switch board receives an average of 4 emergency call in a 10 minute interval. What is the probability that i, there are atmost 2 emergency calls in a 10 minute interval ii, there are exactly 3 emergency calls in a 10 minute interval

Sol: Mean $= \lambda = (4 \text{ calls} / 10 \text{ minutes}) = 4$ calls

$$P(X=x) = P(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-4}4^x}{x!} = \frac{1}{e^4}\frac{4^x}{x!}$$

i, P(atmost 2 calls) $= P(X \le 2)$

$$= P(X=0) + P(X=1) + P(X=2)$$

$$= \frac{1}{e^4} + \frac{1}{e^4}4 + \frac{1}{e^4}\frac{4^2}{2}$$

$$= \frac{1}{e^4}[1+4+8] = \frac{13}{e^4} = 0.23$$

ii, P(Exactly 3 calls)

$$= P(X=3) = 1/e^4 \cdot \frac{4^3}{3!} = \frac{32}{3}e^{-4}$$

$$= 0.19$$

2) Average no of accidents on any day on a national highway is 1.8. Determine the probability that the no of accidents are i, at least one ii, atmost one.

Sol: Mean $\lambda = 1.8$

we have $P(X=x) = P(x) = \frac{e^{-\lambda}\lambda^x}{x!} = e^{-1.8}\frac{(1.8)^x}{x!}$

i, P(atleast one) $= P(x \ge 1) = 1 - P(x=0)$

$$= 1 - \frac{e^{-1.8}(1.8)^0}{0!} = 1 - e^{-1.8}$$

$$= 0.8347$$

ii, P(at most one) $= P(X \le 1) = P(X=0) + P(X=1)$

$$= e^{-1.8}(1 + (1.8)^1)$$

$$= e^{-1.8}(2.8) = 0.46$$

3) If a Bank received on the average 6 bad cheques per day, find the probability that it will receive 4 bad cheques on any given day.

Sol: $P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}$ here $\lambda = 6$

$$\therefore P(X=4) = \frac{e^{-6}6^4}{4!} = \frac{54}{e^6} = 0.1339$$

4) Suppose 2% of the people on the average are left handed. Find i, the probability of finding 3 or more left handed ii, the probability of finding none or one left handed

Sol: Let $x$ be the no. of left handed

Given mean, $\lambda = 2\% = 0.02$

we have $P(x) = \frac{e^{-\lambda}\lambda^x}{x!} = e^{-0.02}\frac{(0.02)^x}{x!}$

i, $P(x \geq 3) = 1 - P(x<3) = 1 - [P(x=0)+P(x=1)+P(x=2)]$

$= 1 - e^{-0.02}\left[1 + 0.02 + \frac{(0.02)^2}{2}\right]$

$= 1 - e^{-0.02}[1.0202]$

$= 1.3077 \times 10^{-6}$

ii, $P(x \leq 1) = P(x=0) + P(x=1)$

$= e^{-0.02}[1 + 0.02] = e^{-0.02}(1.02)$

$= 0.999$

5) If a poisson distribution is such that $P(x=1)\frac{3}{2} = P(x=3)$. Find i, $P(x \geq 1)$ ii, $P(x \leq 3)$

Sol: $\frac{3}{2}P(x=1) = P(x=3)$

$\frac{3}{2}\frac{e^{-\lambda}\lambda}{1!} = \frac{e^{-\lambda}\lambda^{3}}{3!} \Rightarrow \lambda^2 = \frac{3}{2}\times 3!$

$\Rightarrow \lambda^2 = 9 \Rightarrow \lambda = 3$

Hence $P(X=x) = P(x) = \frac{e^{-3}3^x}{x!}$

i, $P(x \geq 1) = 1 - P(x=0) = 1 - \frac{e^{-3}3^0}{0!} = 1 - e^{-3}$

$= 0.950$

ii, $P(x \leq 3) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$

$= e^{-3}\left[\frac{3^0}{0!} + \frac{3}{1!} + \frac{3^2}{2!} + \frac{3^3}{3!}\right]$

$= e^{-3}\left[1 + 3 + \frac{9}{2} + \frac{9}{2}\right] = 13e^{-3} = 0.6472$

6) If a random variable has a poisson distribution such that $P(1) = P(2)$ find i, mean of the distribution ii, $P(4)$ iii, $P(1<x<4)$

Sol: i, $P(x=1) = P(x=2)$

$\frac{e^{-\lambda}\lambda}{1!} = \frac{e^{-\lambda}\lambda^2}{2!} \Rightarrow \lambda = 2$

ii, $P(4) = \frac{e^{-2}2^4}{4!} = \frac{e^{-2}\times 16}{24} = 0.09022$

iii, $P(1<x<4) = P(x=2) + P(x=3)$

$= e^{-2}\left(\frac{2^2}{2!} + \frac{2^3}{3!}\right) = e^{-2}\left(\frac{4}{2} + \frac{8}{6}\right) \Rightarrow 0.4311$

# Unit - III

## Continuous Probability Distribution

When a random variable $x$ takes every value in an interval, it gives rise to continuous distribution of '$x$'. The distribution defined by the variate like temperature, heights and weights are continuous distribution.

## Normal Distribution

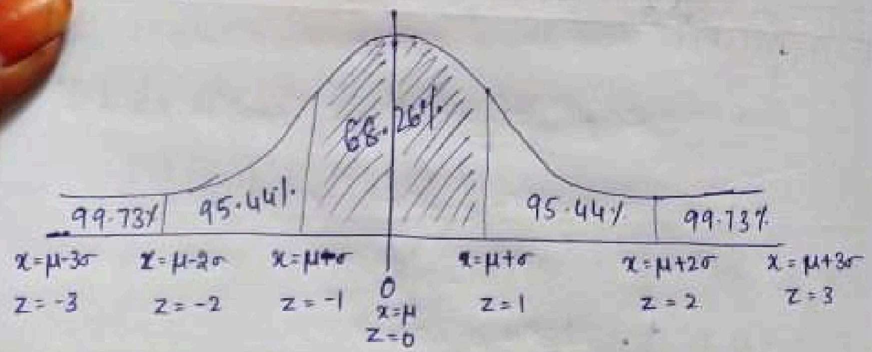
A random variable $x$ is said to have a Normal distribution; If its density function or probability distribution is given by

$$f(x:\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)}{2\sigma^2}} \quad ; -\infty < x < \infty$$
$$-\infty < \mu < \infty, \sigma > 0$$

where $\mu$ = Mean and $\sigma$ = S.d are two parameters of the normal distribution

The random variable $x$ is then said to be a normal random variable or normal variate.

Form of Normal distribution




**How to find probability density of Normal Curve**

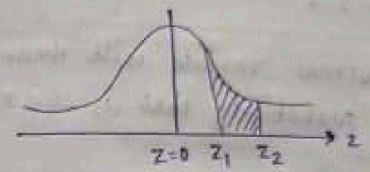
The probability that the normal variate $\bar{x}$ with mean $\mu$ and S.D $\sigma$ lies b/w two specific values $x_1$ and $x_2$ with $x_1 \le x_2$ can be obtained using area under the standard normal curve

**Step 1**: Perform the change of scale $z = \frac{x-\mu}{\sigma}$ and find $z_1$ and $z_2$ corresponding to the values of $x_1$ and $x_2$ respectively

**Step 2**: To find $P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$

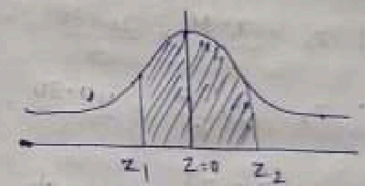
**Case 1**: If both $z_1$ and $z_2$ are positive then

$$P(x_1 \le x \le x_2) = |A(z_2) - A(z_1)|$$



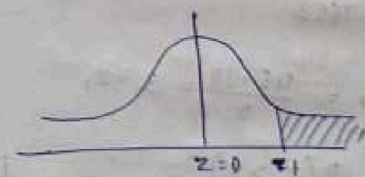

**Case 2**: If both $z_1 < 0$ and $z_2 > 0$ are positive.

$$P(x_1 \le x \le x_2) = A(z_1) + A(z_2)$$



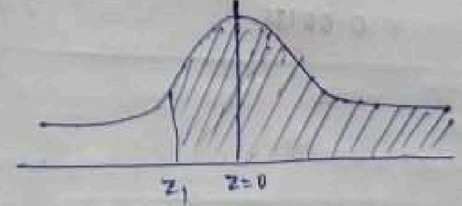

**Case 3**: To find $P(z > z_1)$. If $z_1 > 0$ Then

$$P(z > z_1) = 0.5 - A(z_1)$$



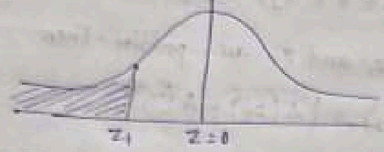

**Case 4**: To find $P(z > z_1)$. If $z_1 < 0$ then

$$P(z > z_1) = 0.5 + A(z_1)$$



**Case 5**: To find $P(z < z_1)$. If $z_1 \leq 0$, then

$$P(z < z_1) = 0.5 - A(z_1)$$



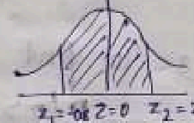

**Eg 1**: If $X$ is a normal variate with mean 30 and s.d '5'. Find the probability that (i) $26 \leq x \leq 60$ (ii) $x \geq 45$

**Sol**: Given mean $\mu = 30$ and s.d, $\sigma = 5$

(i) when $x_1 = 26$; $z_1 = \frac{x_1 - \mu}{\sigma} = \frac{26-30}{5} = -0.8 = z_1$

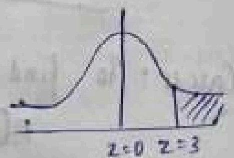
$x_2 = 40$; $z_2 = \frac{x_2 - \mu}{\sigma} = \frac{40-30}{5} = 2 = z_2$

$$P(26 \leq x \leq 40) = P(-0.8 \leq z \leq 2)$$
$$= A(0.8) + A(2)$$
$$= 0.4772 + 0.2881$$
$$= 0.7653$$




(ii) when $x_1 = 45$, $z_1 = \frac{45-30}{5} = 3 = z_1$

$$\therefore P(x \geq 45) = P(z \geq 3)$$
$$= 0.5 - A(3)$$
$$= 0.5 - 0.49865$$
$$= 0.00135$$




**Eg 2**: Suppose the weight of 800 male students are normally distributed with mean $\mu = 140$ pounds and s.d 10 pounds. Find the number of students whose weights are (i) b/w 138 and 148 pounds (ii) more than 152 pounds.

**Sol**: Given $\mu = 140$; $\sigma = 10$

(i) when $x_1 = 138$; $z_1 = \frac{x_1 - \mu}{\sigma} = \frac{138-140}{10} = -0.2 = z_1$

$x_2 = 148$; $z_2 = \frac{148-140}{10} = +0.8 = z_2$

$$P(138 \geq x \geq 148) = P(-0.2 \leq z \leq 0.8)$$
$$= A(z_1) + A(z_2)$$
$$= 0.2881 + 0.0793 = 0.3974$$


No. of students whose weight b/w 138 to 148 $= 0.3974 \times 800 = 294$

(ii) when $x_1 = 152$

$z_1 = \frac{x_1 - \mu}{\sigma} = \frac{152-140}{10} = 1.2 = z_1$

$$P(x > 152) = P(z > 1.2)$$
$$= 0.5 - A(1.2)$$
$$= 0.5 - 0.3849 = 0.1151$$


No. of students whose weight are more than 152 pounds
$= 0.1151 \times 800 = 92$

**Eg 3**: The mean and S.D of the marks obtained by 1000 students in an examination are respectively 34.5 and 16.5. Assuming the normality of the distribution, find the approximate no. of students expected to obtain b/w 30 & 60

**Sol**: Given $\mu = 34.5$ & $\sigma = 16.5$

when $x_1 = 30$; $z_1 = \frac{x_1 - \mu}{\sigma} = \frac{30 - 34.5}{16.5} = -0.27 = z_1$

$x_2 = 60$; $z_2 = \frac{60 - 34.5}{16.5} = 1.54 = z_2$

$P(x_1 \le X \le x_2) = P(z_1 \le z \le z_2)$

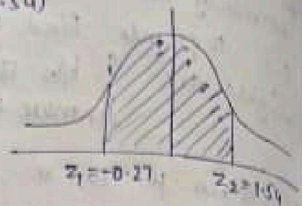
$P(30 \le x \le 60) = P(-0.27 \le z \le 1.54)$

$= A(z_1) + A(z_2)$

$= A(0.27) + A(1.54)$

$= 0.4382 + 0.1084$

$= 0.5916$




$\therefore$ The no. of students who get marks b/w 30 and 60

$= 0.5916 \times 1000 = 591.6$

Hence 592 students get marks b/w 30 and 60

---

Eg 4: In a Normal distribution; 7% of the items are under 35 and 89% are under 63. Determine the mean and variance of the distribution.

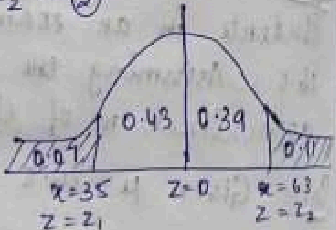
Sol: Let $\mu$ be the mean and $\sigma$ is the S.D of the normal curve 7% of the items are under 35 mean area to the left of the ordinate $x = 35$

Given $P(x < 35) = 0.07$ and $P(x < 63) = 0.89$

$P(x > 63) = 1 - P(x < 63) = 1 - 0.89 = 0.11$

when $x = 35$; $z_1 = \frac{x_1 - \mu}{\sigma} = \frac{35 - \mu}{\sigma} = -z_1$ — ①

$x = 63$; $z_2 = \frac{x_2 - \mu}{\sigma} = \frac{63 - \mu}{\sigma} = z_2$ — ②




From the figure, we have

$P(0 < z < z_2) = 0.39 \Rightarrow z_2 = 1.23$

and $P(0 < z < z_1) = 0.43 \Rightarrow z_1 = 1.48$ } From table

From equ ① $\frac{35 - \mu}{\sigma} = -1.48$

$35 = -1.48\sigma + \mu$ — ③

From equ ② $\frac{63 - \mu}{\sigma} = 1.28$

$63 = 1.28\sigma + \mu$ — ④

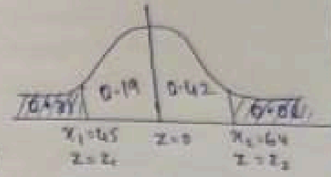
From equ ③ & ④ $\boxed{\mu = 50.3,\ \sigma^2 = 106.75}$

Eg 5: In a normal distribution 31% of the items are under 45 and 8% are over 64. Find mean and variance

Sol: $x_1 = 45$ $\quad x_2 = 64$

$z_1 = \frac{x_1 - \mu}{\sigma}$ $\quad z_2 = \frac{x_2 - \mu}{\sigma}$

$z_1 = \frac{45 - \mu}{\sigma}$ — ① $\quad z_2 = \frac{64 - \mu}{\sigma}$ — ②




$P(0 < z < z_1) = A(z_1) = 0.19 \Rightarrow z_1 = -0.5$

$P(0 < z < z_2) = A(z_2) = 0.42 \Rightarrow z_2 = 1.4$ } from table

From eq ①

$-0.5 = \frac{45 - \mu}{\sigma}$

$-0.5\sigma = 45 - \mu$

$\mu - 0.5\sigma = 45$ — ③

From eq ②

$1.4 = \frac{64 - \mu}{\sigma}$

$1.4\sigma = 64 - \mu$

$\mu + 1.4\sigma = 64$ — ④

From ③ & ④

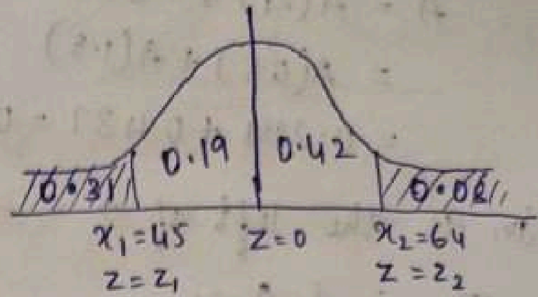
$\boxed{\mu = 50,\ \sigma = 10}$

Eg5: ... are under 45 and 8% are over 64. Find mean and variance.

Sol: $x_1 = 45 \qquad x_2 = 64$

$z_1 = \dfrac{x_1 - \mu}{\sigma} \qquad z_2 = \dfrac{x_2 - \mu}{\sigma}$

$z_1 = \dfrac{45 - \mu}{\sigma} \quad — (1) \qquad z_2 = \dfrac{64 - \mu}{\sigma} \quad — (2)$




$P(0 < z < z_1) = A(z_1) = 0.19 \Rightarrow z_1 = -0.5$
$P(0 < z < z_2) = A(z_2) = 0.42 \Rightarrow z_2 = 1.4$ } from table

From eq (1)

$-0.5 = \dfrac{45 - \mu}{\sigma}$

$-0.5\sigma = 45 - \mu$

$\mu - 0.5\sigma = 45 \quad — (3)$

From eq (2)

$1.4 = \dfrac{64 - \mu}{\sigma}$

$1.4\sigma = 64 - \mu$

$\mu + 1.4\sigma = 64 \quad — (4)$

From (3) & (4)

$$\boxed{\mu = 50 \;, \quad \sigma = 10}$$

---

Eg6: If $\bar{x}$ is a normal variate, Find the area A

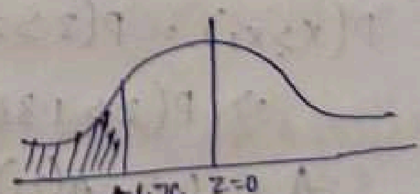
i, To the left of $z = -1.78$ ii, To the right of $z = -1.45$

iii, Corresponding to $-0.8 \le z \le 1.53$

iv, To the left of $z = -2.52$ and to the right of $z = 1.83$

Sol: i, To the left of $z = -1.78$

Required area, $A = 0.5 - A(z_1)$




$A = 0.5 - A(1.78)$

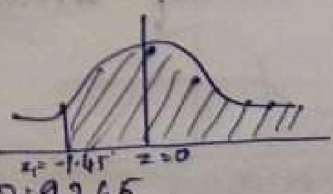
$= 0.5 - 0.4625 = 0.0375$

ii, To the right of $z = -1.45$

Required Area, $A = 0.5 + A(z_1)$

$= 0.5 + A(1.45)$

$= 0.5 + 0.4265 = 0.9265$



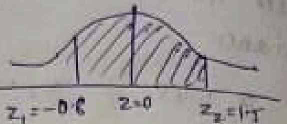
iii, Corresponding to $-0.8 \le z \le 1.5$

Required area

$A = A(z_1) + A(z_2)$

$= A(0.8) + A(1.5)$

$= 0.2881 + 0.4332 = 0.7251$




iv, To the left of $z = -2.52$ and to the right of $z = 1.83$

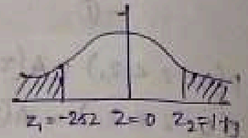
Required Area

$A = [0.5 - A(z_1)] + [0.5 - A(z_2)]$

$A = [0.5 - A(-2.52)] + [0.5 - A(1.83)]$

$A = (0.5 - 0.4941) + (0.5 - 0.4664)$

$A = 0.0395$




---

Eg7: If the masses of 300 students are normally distributed with mean 68 kgs and S.D 3 kgs, How many students have masses i, Greater than 72 kg ii, less than or equal to 64 kg.

Sol: Given $\mu = 68$, $\sigma = 3$

i, When $x_1 = 72 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{72-68}{3}$

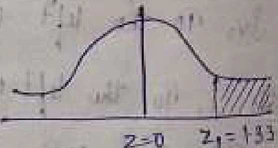
$z_1 = 1.33$

$P(x > x_1) = P(z > z_1) = P(x > 72)$

$\Rightarrow P(z > 1.33)$

$A = 0.5 - A(z_1) = 0.5 - A(1.33)$

$= 0.5 - 0.4082 = 0.0918$




No of students with more than 72 kgs.

$= 300 \times 0.0918$

$= 28$.

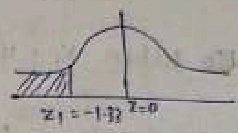
ii, When $x_1 = 64 \Rightarrow z_1 = \frac{x_1 - \mu}{\sigma} = \frac{64-68}{3}$.

$z_1 = -1.33$

$P(x \le x_1) = P(z \le z_1)$

$A = 0.5 - A(z_1)$

$A = 0.5 - 0.4082 = 0.0918$




No of students have masses less than or equal to 64 kg $= 0.0918 \times 300 = 28$

---

Eg8: In a sample of 1000 cases, the mean of a certain test is 14 and S.D is 2.5 Assuming the distribution to be normal, find

i, How many students score b/w 12 and 15

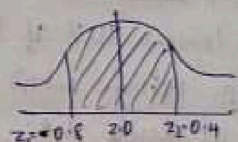
ii, How many score above 18? iii, How many score below 16?

Sol: Given $\mu = 14$, $\sigma = 2.5$

Let the variable $\bar{x}$ denote the score in the test

i, When $x_1 = 12$ ; $z_1 = \frac{12-14}{2.5} = -0.8$

When $x_2 = 15$ ; $z_1 = \frac{15-14}{2.5} = 0.4$




Required Area $= P(x_1 \le x \le x_2) = P(z_1 \le z \le z_2)$

$= A(z_1) + A(z_2)$

$= A(0.8) + A(0.4) = 0.2881 + 0.1554 = 0.4435$

No. of students score b/w 12 and 15 is $0.4435 \times 1000 = 443$

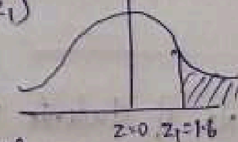
ii, When $x_1 = 18 \Rightarrow z_1 = \frac{18-14}{2.5} = 1.6$

Required area $= P(x > x_1) = P(z > z_1)$

$= 0.5 - A(z_1)$

$= 0.5 - A(1.6)$

$= 0.5 - 0.4452 = 0.0548$



∴ No. of students score above 18 is
$0.0548 \times 1000 = 54.8 \simeq 55$

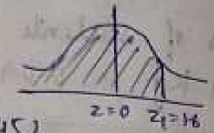
iii, When $x_1 = 18$, $z_1 = \frac{18-14}{2.5} = 1.6$

Required area $= P(x < x_1) = P(z < z_1)$

$= 0.5 + A(z_1)$

$= 0.5 + A(1.6)$

$= 0.5 + 0.4452 = 0.9452$




No of students score below 18 is
$0.9452 \times 1000 = 945$

---

Eg 9: The marks obtained in Mathematics by 1000 students is normally distributed with mean 78% and S.D 11%, Determine:
i, How many students get marks about 90%.
ii, What was the highest mark obtained by the lowest 10% of the students
iii, Within what limits did middle of 90% of the students lie.

Sol: Given mean $\mu = 0.78$, $\sigma = 0.11$

i, When $x = 0.9$, $z = \frac{x-\mu}{\sigma} = \frac{0.9-0.78}{0.11} = 1.09 = z_1$

[To find how many students got above 90%, we need to first find $P(x > 90\%) = P(x > 0.9)$]

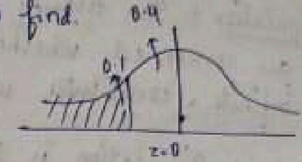
$\Rightarrow P(x > 0.9) = P(z > 1.09)$

$= 0.5 - A(1.09)$

$= 0.5 - 0.3621 = 0.1379$


No. of Students who got more than 90% $= 1000 \times 0.1379$
$= 137.9 \Rightarrow 138$ students

ii, To find highest mark obtained by lowest 10% of students. i.e., 0.1 area to the left to z
For the 0.4 area, we need to find corresponding z value
i.e., $0.5 - 0.1 = 0.4$




So, for 0.4, $-1.28$ (from table)

for $z_1 = -1.28 = \frac{x-\mu}{\sigma} = \frac{x-0.78}{0.11} \Rightarrow (-1.28)(0.11) = x - 0.78$
$\Rightarrow x = 0.78 - (1.28)(0.11)$
$\Rightarrow x = 0.6392$

Hence, the highest mark obtained by lowest 10% of students $= 0.6392 \times 1000 = 64\%$

iii, Within what limits did the middle of 90% of students lie. Middle 90% are i.e 0.9 area is distributed equally in the curve. for 0.45 area corresponding values of $z = \frac{1.64 + 1.65}{2} = 1.645$


Left of $z = 0$, $z_2 = -1.645$, Right of $z = 0$, $z_1 = 1.645$

∴ We need to find the marks (i.e x value)

When $z_1 = 1.64$, $z_1 = \frac{x_1 - \mu}{\sigma} \Rightarrow x_1 = z_1\sigma + \mu$
$x_1 = (1.64)(0.11) + 0.78$
$x_1 = 0.9604$ (or) 96.04%

$z_2 = -1.64$, $z_2 = \frac{x_2 - \mu}{\sigma} \Rightarrow x_2 = z_2\sigma + \mu$
$x_2 = (-1.64)(0.11) + 0.78$
$x_2 = 0.5996$ (or) 59.96%

∴ Thus middle 90% have marks in b/w 60 to 90%

# Sampling Distribution

**Population:** The totality of observations with which we are concerned, whether this number be finite or infinite, constitutes what we call population

**Defn:** Population is the aggregate or totality of statistical data forming a subject of investigation

Eg: ① The population of the heights of Indians
② The population of Nationalised Banks in India, etc....

→ The number of observations in the population is defined to be the size of the population. It may be finite or infinite. Size of the population is denoted by **N**

**Sample:** A Sample is a subset of population and the no. of objects in the sample is called the size of the sample. It is denoted by **n**.

The process of selection of sample is called sampling

→ If N is the size of the population and n is the sample size, then

i, the number of samples with replacement = $N^n$

ii, the no. of samples without replacement = $^NC_n$

**Large Sample:** If the size of the sample $(n) \geq 30$, the sample is said to be large sample.

**Small Sample:** If the size of the sample $(n) < 30$, the sample is said to be small sample (or) exact sample

**Statistics:** Measures obtained from the sample of the population are called statistics

Eg: Sample mean, sample variance are called statistics

**Parameters:** The measures obtained from the population are called parameters

Eg. Population mean, population variance

**Sample Mean**

If $x_1, x_2, \ldots x_n$ represent a random sample of size n, then the sample mean is defined by the statistic

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$$

Sample mean is denoted by $\bar{x}$ or $\bar{x}$

∴ Sample Variance: $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$

Sample standard deviation: $SD = S = \sqrt{S^2} = \sqrt{\text{Sample Variance}}$

**Central limit theorem:**

If $\bar{x}$ be the mean of a sample size n drawn from a population with mean $\mu$ and S.D $\sigma$ then the standardized sample mean.

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

is a random variable whose distribution function approaches that of the standard normal distribution $N(z; 0, 1)$ as $n \to \infty$

**Infinite Population:** Suppose the samples are drawn from an infinite population i.e $N \to \infty$ (or) Sampling is done with replacement, then

i, Mean of sampling distribution of means $\mu_{\bar{x}} = \mu$

i.e $E(\bar{x}) = \mu$

ii, $Var(\bar{x}) = \frac{\sigma^2}{n}$ (or) $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$

$\sigma^2_{\bar{x}} \to$ Variance of sampling distribution of means

**Finite Population:** Consider a finite population of size N with mean $\mu$ and standard deviation $\sigma$. Draw all possible samples of size n without replacement from this population. Then

i, The mean of the sampling distribution of means is $\mu_{\bar{x}} = \mu$ i.e $E(\bar{x}) = \mu$

ii, Variance is $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ & S.D $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

$\frac{N-n}{N-1}$ is called finite population correction factor

1) Eg: What is the value of correction factor if $n = 5$ & $N = 200$

Sol: Correction factor $= \frac{N-n}{N-1} = \frac{200-5}{200-1} = \frac{195}{199}$

$= 0.98$

2) How many different samples of size two can be chosen, from a finite population of size 25

Sol. No. of samples $= {}^{25}C_2$

3) Find the correction factor for $n = 10$ & $N = 100$

Sol. Correction factor $= \frac{N-n}{N-1} = \frac{100-10}{100-1} = \frac{90}{99} = 0.9090$

**Sampling Distribution Problems**

(a) A population consists of 5 numbers 2, 3, 6, 8, 11. Consider all samples of size two which can be drawn with replacement from the population. Find

(a) The population mean

(b) Variance of population

(c) The mean of sampling distribution of means

(d) Variance of sampling distribution of means.

Sol: Population $= \{2, 3, 6, 8, 11\}$

(a) Mean of the population $= \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6 = \mu$

(b) Variance of the population $\sigma^2$ is given by

$$\sigma^2 = \sum \frac{(x_i - \bar{x})^2}{n} \text{ (or) } \sum \frac{(x_i - \mu)^2}{n}$$

$$\sigma^2 = \frac{(2-6)^2 + (3-6)^2 + (6-6)^2 + (8-6)^2 + (11-6)^2}{5} = \frac{16+9+4+25}{5}$$

$$\sigma^2 = \frac{54}{5} = 10.8 \Rightarrow \sigma = \sqrt{10.8} = 3.29$$

(c) Now samples have to be drawn from population of size 2 with replacement. No. of samples $= N^n$

$= 5^2 = 25$

where N = population size, n = sample size

(2,2) (2,3) (2,6) (2,8) (2,11)
(3,2) (3,3) (3,6) (3,8) (3,11)
(6,2) (6,3) (6,6) (6,8) (6,11)
(8,2) (8,3) (8,6) (8,8) (8,11)
(11,2) (11,3) (11,6) (11,8) (11,11)

Now we compute mean of each of the sample. The set of 25 means $\bar{x}$ of the above 25 samples, gives rise to distribution of means of the samples known as <u>sampling distribution of means</u>

Sample means are

$$\left\{\begin{array}{lllll} \frac{2+2}{2}=2 & \frac{2+3}{2}=2.5 & \frac{2+6}{2}=4 & \frac{2+8}{2}=5 & \frac{2+11}{2}=6.5 \\ \frac{3+2}{2}=2.5 & \frac{3+3}{2}=3 & \frac{3+6}{2}=4.5 & \frac{3+8}{2}=5.5 & \frac{3+11}{2}=7 \\ \frac{6+2}{2}=4 & \frac{6+3}{2}=4.5 & \frac{6+6}{2}=6 & \frac{6+8}{2}=7 & \frac{6+11}{2}=8.5 \\ \frac{8+2}{2}=5 & \frac{8+3}{2}=5.5 & \frac{8+6}{2}=7 & \frac{8+8}{2}=8 & \frac{8+11}{2}=9.5 \\ \frac{11+2}{2}=6.5 & \frac{11+3}{2}=7 & \frac{11+6}{2}=8.5 & \frac{11+8}{2}=9.5 & \frac{11+11}{2}=11 \end{array}\right\} \longrightarrow (1)$$

Mean of the sampling distribution of means $= \mu_{\bar{x}}$

$$= \frac{2+2.5+4+5+6.5+2.5+3+4.5+5.5+\dots+6.5+7+8.5+9.5+11}{25}$$

$$\mu_{\bar{x}} = \frac{150}{25} = 6 = \mu \qquad \boxed{\therefore \mu_{\bar{x}} = \mu}$$

(2) Variance $\sigma_{\bar{x}}^2$ of the sampling distribution of means is obtained by subtracting mean 6 from each number in (1) & squaring the result, adding all 25 numbers thus obtained, & dividing by 25.

$$\sigma_{\bar{x}}^2 = \frac{\sum(\bar{x}_i - \mu_{\bar{x}})^2}{n}$$

$$= \frac{(2-6)^2+(2.5-6)^2+(4-6)^2+(5-6)^2+(6.5-6)^2+(2.5-6)^2+(3-6)^2+(4.5-6)^2+\dots+(6.5-6)^2+(7-6)^2+(8.5-6)^2+(9.5-6)^2+(11-6)^2}{25}$$

$$\sigma_{\bar{x}}^2 = \frac{135}{25} = 5.40 \Rightarrow \sigma_{\bar{x}} = \sqrt{5.4} = 2.32$$

for infinite population, $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \Rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$$\sigma_{\bar{x}} = \frac{3.29}{\sqrt{2}} = 2.32$$

<u>Same example</u> <u>without</u> <u>replacement</u>

Total no. of samples without replacement is ${}^{N}C_{n} = {}^{5}C_{2} = 10$ samples of size 2

The 10 samples are:

(2,3) (2,6) (2,8) (2,11)
(3,6) (3,8) (3,11)
(6,8) (6,11)
(8,11)

The corresponding sample means are:

$$\left\{\begin{array}{llll} \frac{2+3}{2}=2.5 & \frac{2+6}{2}=4 & \frac{2+8}{2}=5 & \frac{2+11}{2}=6.5 \\ \frac{3+6}{2}=4.5 & \frac{3+8}{2}=5.5 & \frac{3+11}{2}=7 & \\ \frac{6+8}{2}=7 & \frac{6+11}{2}=8.5 & & \\ \frac{8+11}{2}=9.5 & & & \end{array}\right.$$

Mean of sampling distribution of means:

$$\mu_{\bar{x}} = \frac{2.5+4+5+6.5+4.5+5.5+7+7+8.5+9.5}{10} = 6$$

$$\mu_{\bar{x}} = \mu$$

Variance of sampling distribution of means

$$\sigma_{\bar{x}}^2 = \frac{(2.5-6)^2+(4-6)^2+(6.5-6)^2+\dots+(7-6)^2+(8.5-6)^2+(9.5-6)^2}{10}$$

$$= 4.05$$

$$\Rightarrow \sigma_{\bar{x}} = 2.01$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right) = \frac{10.8}{2}\left(\frac{5-2}{5-1}\right) = 4.05$$ for sampling without replacement

② A population consists of six numbers 4, 8, 12, 16, 20, 24. Consider all samples of size two which can be drawn without replacement from this population. Find

ⓐ The population mean
ⓑ The population standard deviation
ⓒ The mean of the sampling distribution of means
ⓓ The standard deviation of the sampling distribution of means

Sol. ⓐ The mean of the population $= \frac{4+8+12+16+20+24}{6}$

$= \frac{84}{6} = 14$

ⓑ Variance of the population $= \sigma^2 = \frac{1}{N}\sum(x_i-\bar{x})^2$

$= \frac{1}{6}\left[(4-14)^2+(8-14)^2+(12-14)^2+(16-14)^2+(20-14)^2+(24-14)^2\right]$

$= \frac{1}{6}\left[100+36+4+4+36+100\right] = \frac{280}{6} = 46.67$

$SD = \sigma = \sqrt{46.67} = 6.83$

ⓒ No. of sample of size 2 that can be drawn from population without replacement : ${}^6C_2 = 15$

(4,8) (4,12) (4,16) (4,20) (4,24)
(8,12) (8,16) (8,20) (8,24)
(12,16) (12,20) (12,24)
(16,20) (16,24)
(20,24)

The means are :
6 8 10 12 14
10 12 14 16
14 16 18
18 20
22

Mean of the sampling distribution of means is

$\mu_{\bar{x}} = \frac{6+8+10+12+14+10+12+14+16+14+16+18+18+20+22}{15}$

$\mu_{\bar{x}} = \frac{210}{15} = 14$

ⓓ Variance of sampling distribution of means $= \sigma^2_{\bar{x}}$

$= \frac{1}{15}\sum(x_i-\mu_{\bar{x}})^2$

$= \frac{1}{15}\left[(6-14)^2+(8-14)^2+(10-14)^2+(12-14)^2+(14-14)^2+(10-14)^2+(12-14)^2 + (14-14)^2+(16-14)^2+\cdots+(20-14)^2+(22-14)^2\right]$

$= \frac{1}{15}\left[64+36+16+4+0+16+4+0+4+0+4+16+16+36+64\right]$

$= \frac{280}{15} = 18.67$

$\sigma_{\bar{x}} = \sqrt{18.67} = 4.32\sqrt{2}$

---

③ If the population is 3, 6, 9, 15, 27

ⓐ List all possible samples of size 3 that can be taken without replacement from the finite population
ⓑ Calculate the mean of each of the sampling distribution of means.
ⓒ Find the standard deviation of sampling distribution of means.

Sol: Mean of the population $= \frac{3+6+09+15+27}{5} = \frac{60}{5} = 12$

S.D of population $= \sigma = \sqrt{\frac{(3-12)^2+(6-12)^2+(9-12)^2+(15-12)^2+(27-12)^2}{5}}$

$= \sqrt{\frac{81+36+9+9+225}{5}} = \cancel{\sqrt{\frac{400}{5}} = \sqrt{80} = 4\sqrt{5}} = 8.9442$

$= \sqrt{\frac{360}{5}} = 8.4853$

Total no. of samples without replacement $= {}^{N}C_{n} = {}^{5}C_{3} = 10$

The 10 samples are

(3,6,9) (3,6,15) (3,9,15) (3,6,27) (3,9,27) (3,15,27) (6,9,15)
(6,9,27) (6,15,27) (9,15,27)

Mean of each of the samples:

6, 8, 9, 12, 13, 15, 10, 12, 16, 17

Mean of the sampling distribution of means is

$$\mu_{\bar{x}} = \frac{6+8+9+12+13+15+10+12+16+17}{10} = \frac{120}{10} = 12$$

$$\sigma_{\bar{x}}^2 = \frac{1}{10}\left[(6-12)^2+(8-12)^2+(9-12)^2+(12-12)^2+(13-12)^2+(15-12)^2 + (10-12)^2+(16-12)^2+(17-12)^2\right]$$

$$= \frac{120}{10} = 12 \Rightarrow \sigma_{\bar{x}} = \sqrt{12} = 3.464$$

---

(4) Let a population consist of the four numbers 1, 5, 6, 8. Consider all possible samples of size two that can be drawn without replacement from this population.

i, the population mean

ii, the population standard deviation.

iii, The mean of the sampling distribution of means

iv, The standard deviation of the sampling distribution of means.

Sol: i, Population mean $= \frac{1+5+6+8}{4} = \frac{20}{4} = 5$

ii, S.D of population $= \sigma = \sqrt{\frac{(1-5)^2+(5-5)^2+(6-5)^2+(8-5)^2}{4}}$

$$= \sqrt{\frac{16+0+1+9}{4}} = \sqrt{\frac{25}{4}} = \frac{5}{2} = 2.5$$

iii, Total no. of samples without replacement $= {}^{N}C_{n} = {}^{4}C_{2} = 6$

The 6 samples are:

(1,5) (1,6) (1,8)
(5,6) (5,8)
(6,8)

Mean of each of the samples

3, 3.5, 4.5, 5.5, 6.5, 7

Mean of the sampling distribution of means is

$$\mu_{\bar{x}} = \frac{3+3.5+4.5+5.5+6.5+7}{6} = 5$$

Variance of the sampling distribution of means is

$$\sigma_{\bar{x}}^2 = \frac{1}{6}\left[(3-5)^2+(3.5-5)^2+(4.5-5)^2+(5.5-5)^2+(6.5-5)^2+(7-5)^2\right]$$

$$= \frac{1}{6}\left[4+2.25+0.25+0.25+2.25+4\right]$$

$$\sigma_{\bar{x}}^2 = \frac{13}{6} = 2.16666 \Rightarrow \sigma_{\bar{x}} = \sqrt{2.16666} = 1.471$$

---

(5) Take population 1, 2, 3, 4, 5, 6. Sample of size 2 with replacement.

i, the population mean

ii, the population standard deviation

iii, The mean of the sampling distribution of mean

iv, The standard deviation of sampling distribution of means.

Sol: 1, 2, 3, 4, 5, 6

i, Population mean $= \frac{1+2+3+4+5+6}{6} = 3.5$

ii, Variance $\sigma^2 = \frac{1}{N}\sum(x_i - \bar{x})^2$

$= \frac{1}{6}\left[(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2\right]$

$= \frac{1}{6}\left[6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.25\right]$

$\sigma^2 = \frac{35}{12} = 2.917 \Rightarrow \sigma = \sqrt{2.197} = 1.707$

iii, No. of samples of size 2 that can be drawn with replacement is $N^n = 6^2 = 36$

(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)

The means are

| | | | | | |
|---|---|---|---|---|---|
| 1, | 1.5 | 2 | 2.5 | 3 | 3.5 |
| 1.5 | 2 | 2.5 | 3 | 3.5 | 4 |
| 2 | 2.5 | 3 | 3.5 | 4 | 4.5 |
| 2.5 | 3 | 3.5 | 4 | 4.5 | 5 |
| 3 | 3.5 | 4 | 4.5 | 5 | 5.5 |
| 3.5 | 4 | 4.5 | 5 | 5.5 | 6 |

$\mu_{\bar{x}} = \frac{1}{36}\left[1+1.5+2+2.5+3+3.5+ \dots + 3.5+4+4.5+5+5.5+6\right]$

$= \frac{126}{36} = 3.5$

iv, Variance $\sigma^2_{\bar{x}} = \frac{\left[(1-3.5)^2 + (1.5-3.5)^2 + (2-3.5)^2 + \dots + (6-3.5)^2\right]}{36}$

$= \frac{52.5}{36} = 1.46$

SD $\sigma_{\bar{x}} = \sqrt{1.46} = 1.208$

1) A random sample of size 100 is taken from an infinite population having the mean $\mu = 76$ and the variance $\sigma^2 = 256$. What is the probability that $\bar{x}$ will be between 75 and 78.

Sol: Given sample size $n = 100$, $\mu = 76$, $\sigma^2 = 256 \Rightarrow \sigma = 16$

Since n is large, sample mean $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{\sqrt{n}}\right)$

$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

To find $P(75 \le \bar{x} \le 78)$.

Let $\bar{x}_1 = 75$, $z_1 = \frac{\bar{x}_1 - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{75-76}{\frac{16}{\sqrt{100}}} = \frac{-1(10)}{16} = -0.625$

$\bar{x}_2 = 78$, $z_2 = \frac{\bar{x}_2 - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{78-76}{\frac{16}{\sqrt{100}}} = \frac{2(10)}{6} = 1.25$

$\therefore P(75 \le \bar{x} \le 78) = P(z_1 \le z \le z_2)$

$= A(0.625) + A(1.25)$

$= 0.2334 + 0.3944 = 0.628$

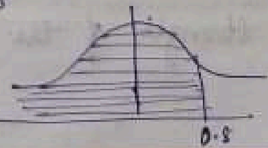
2) The mean height of students in a college is 155 cms and standard deviation is 15. What is the ~~prop~~ probability that the mean height of 36 students is less than 157 cms

Sol: Given $\mu$ = mean height of students of a college = 155 cms

$\sigma = 15$, $n = 36$ = sample size.

To find $P(\bar{x} < 157)$:

$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{157-155}{\frac{15}{\sqrt{36}}} = \frac{2 \times 6}{15} = 0.8$



$P(\bar{x} < 157) = P(z < 0.8) = 0.5 + P(0 < z < 0.8)$

$= 0.5 + A(0.8)$

$= 0.5 + 0.2881$

$= 0.7881$

$\therefore$ Probability that the mean height of 36 students is less than 157 cms $= 0.7881$

repeated

3) A random sample of size 100 is taken from an infinite population having the mean $\mu = 76$ and the variance $\sigma^2 = 256$. What is the probability that $\bar{x}$ will be between 75 and 78.

Sol: $n = 100$ Given $\mu = 76$ $\sigma^2 = 256$

$\therefore \sigma = 16$

To find $P(75 \le \bar{x} \le 78)$:

Let ~~first~~ ~~P~~ $\bar{x}_1 = 75$, $z_1 = \dfrac{\bar{x}_1 - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{75 - 76}{\frac{16}{\sqrt{100}}} = \dfrac{-10}{16} = -0.625$

$\bar{x}_2 = 78$, $z_2 = \dfrac{\bar{x}_2 - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{78 - 76}{\frac{16}{\sqrt{100}}} = \dfrac{2(10)}{16} = 1.25$

$\therefore P(75 \le \bar{x} \le 78) = P(z_1 \le z \le z_2)$

$= P(-0.625 \le z \le 1.25)$

$= A(0.625) + A(1.25)$

$= 0.2334 + 0.3944 = 0.628$

3) If the mean of breaking strength of copper wire is 575 lbs, with a SD of 8.3 lbs. How large a sample must be used in order that there will be one chance in 100 that the mean breaking strength of the sample is less than 572 lbs?

Sol: Given mean of strength of copper wire = 575 lbs

$\mu = 575$ lbs, $\sigma = 8.3$ lbs

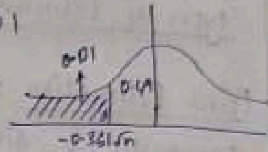
Given that $P(\bar{x} < 572)$ is $\frac{1}{100} = 0.01$

when $\bar{x} = 572$, $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{572 - 575}{\frac{8.3}{\sqrt{n}}} = \dfrac{-3\sqrt{n}}{8.3}$

$z = 0.361\sqrt{n}$

$P(\bar{x} < 572) = P(z < -0.361\sqrt{n}) = 0.01$

$= P(z > 0.361\sqrt{n}) = 0.01$




when area is 0.49, $z = 2.33$

$z = 2.33 = 0.361\sqrt{n}$

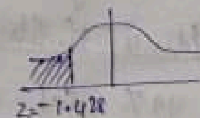
$\sqrt{n} = \dfrac{2.33}{0.36} \Rightarrow n = \left(\dfrac{2.33}{0.36}\right)^2 = 41.6 \simeq 42$

4) A normal population has a mean of 0.1 and standard deviation of 2.1. Find the probability that mean of a sample of size 900 will be negative.

Sol: Given $\mu = 0.1$, $\sigma = 2.1$

To find $P(\bar{x} < 0)$:

$z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{0 - 0.1}{\frac{2.1}{\sqrt{900}}} = \dfrac{(-0.1)}{2.1}(30) = -1.428$




$\therefore P(\bar{x} < 0) = P(z < -1.428)$

$= 0.5 - P(z < 1.428)$

$= 0.5 - 0.4222$

$= 0.0778$

# Estimation

**Estimate**: An estimate is a statement made to find an unknown population parameter.

**Estimator**: The procedure or rule to determine an unknown population parameter is called an Estimator.

**Types of Estimation**
- Point Estimation
- Interval Estimation

**Point Estimation**: If an estimate of the population parameter is given by a single value, then the estimate is called a point Estimation of the parameter.

Eg: 70, 80, 90 marks

**Interval Estimation**: If an estimate of a population parameter is given by two different values b/w which parameter may be consider to lie, then the estimate is called an Interval estimation of the parameter.

Eg: 20-30, 70-80 marks

**Confidence limits for population mean '$\mu$'**

i, 99% C.L are $\bar{x} \pm 2.58$ (S.E of $\bar{x}$) ($z_{\alpha/2}$)

ii, 95% C.L are $\bar{x} \pm 1.96$ (S.E of $\bar{x}$) ($z_{\alpha/2}$)

iii, 90% C.L are $\bar{x} \pm 1.64$ (SE of $\bar{x}$) ($z_{\alpha/2}$)

**Maximum Error of Estimate E**

The max Error of E with $(1-\alpha)$ probability is

$$E_{max} = z_{\alpha/2}\left[\frac{\sigma}{\sqrt{n}}\right]$$

1) What is in size of the smallest sample required to estimate an unknown proportion to within a maximum error of 0.06 with atleast 95% confidence

Sol: The max Error E = 0.06

Confidence limit $z_{\alpha/2} = 1.96$ at 95%.

Here P is not given, so we take $P = 1/2$, $q = 1/2$

$$\sigma = \sqrt{pq} = \sqrt{\frac{1}{2}\times\frac{1}{2}} = \frac{1}{2}$$

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{z_{\alpha/2}\,\sigma}{E} \Rightarrow n = \left(\frac{1.96}{0.06}\right)^2 \times \frac{1}{4}$$

$$n = 266.78 \approx 267 \qquad \boxed{n = 267}$$

2) If we can assert with 95% the maximum error is 0.05 and P = 0.2 find the size of the sample.

Sol: $P = 0.2$ $\quad q = 0.8$

$z_{\alpha/2} = 1.96$ at 95% (C.S)

$E = 0.05$

$$\sigma = \sqrt{(0.2)(0.8)}$$
$$\sigma = \sqrt{0.16} = 0.4$$

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$n = \left(\frac{1.96}{0.05}\right) \times 0.4 = 245.86 \simeq 246$$

$$\boxed{n = 246}$$

3) Assuming that $\sigma = 20.0$, how large a random sample be taken to assert with probability 0.95 that the sample mean will not differ from the true mean by more than 3.0 points.

Sol: Given maximum error $E = 3.0$, $\sigma = 20.0$

we have $z_{\alpha/2} = 1.96$ at 95%

$$n = \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2 = \left(\frac{1.96 \times 20}{3}\right)^2 = 170.74$$

$$\boxed{n = 171}$$

---

4) It is desired to estimate the mean no. of hours of continuous use until a certain computer will first require repairs. If it can be assumed that $\sigma = 48H$. Then how large a sample be needed so that one will be able to assert with 90% confidence that the sample mean is off by at most 10 H

Sol: Given, $\sigma = 48H$, $E = 10H$, $z_{\alpha/2} = 1.64$ at 90%.

$$n = \left(\frac{z_{\alpha/2}\,\sigma}{E}\right)^2 = \left(\frac{1.64 \times 48}{10}\right)^2$$

$$n = 61.96 \simeq 62$$

$$\boxed{n = 62}$$

---

Confidence Interval for $\mu$, $\sigma$ known.

If $\bar{x}$ is the mean of a random sample of size $n$ from the population with known variance $\sigma^2$; $(1-\alpha)100\%$ confidence interval for $\mu$ is given by

$$C.I = \left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

1) Determine a 95% confidence interval for the mean of a normal distribution with variance 0.25 using of $n = 100$ values with mean 212.3

Sol: We have $n = 100$; $\bar{x} = 212.3$, SD $\sigma = \sqrt{0.25}$ and $z_{\alpha/2} = 1.96$ (for 95%)

we know that 95% confidence interval is

$$C.I = \left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$C.I = \left(212.3 - 1.96\sqrt{\frac{0.25}{100}},\ 212.3 + 1.96\sqrt{\frac{0.25}{100}}\right)$$

$$C.I = (212.3 - 0.098,\ 212.3 + 0.098)$$

$$C.I = (212.202,\ 212.398)$$

---

2) Find 95% C.I for the mean of a normally distributed population from which the following sample was taken 15, 17, 10, 18, 16, 9, 7, 11, 13, 14

Sol: We have $\bar{x} = \dfrac{15+17+10+18+16+9+7+11+13+14}{10}$

$$\bar{x} = \frac{130}{10} = 13$$

$$s^2 = \sigma^2 = \frac{1}{9}\left[(15-13)^2 + (17-13)^2 + (10-13)^2 + \cdots + (13-13)^2 + (14-13)^2\right]$$

$\sigma^2 = \frac{40}{3}$ , $\sigma = \sqrt{\frac{40}{3}}$ , $n = 10$ , $Z_{\alpha/2} = 1.96$

$\bar{x} = 13$

$CI = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} , \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

$CI = \left(13 - 1.96\frac{\sqrt{40}}{\sqrt{30}} , 13 + 1.96\sqrt{\frac{40}{30}}\right)$

$CI = (10.74 , 15.26)$

---

3) Measurements of the weights of a random sample of 200 ball bearing made by a certain machine during one week showed a mean of 0.824 and a S.D of 0.042. Find maximum error at 95% CI

Sol: $\bar{x} = 0.824$ , $\sigma = 0.042$ , $Z_{\alpha/2} = 1.96$ at 95%

$n = 200$

$$E = Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \frac{1.96 \times 0.042}{\sqrt{200}} = 0.0058$$

$C.I = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} , \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

$C.I = (0.824 - 0.0058 , 0.824 + 0.0058)$

$C.I = (0.8182 , 0.8298)$

---

6) A sample of size 300 was taken whose variance is 225 and mean 54. Have 95% CI for the mean

Sol: Given $\bar{x} = 54$ , $\sigma^2 = 225 \Rightarrow \sigma = 15$ , $n = 300$.

$Z_{\alpha/2} = 1.96$ at 95%

$$E = Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}} = 1.96\frac{15}{\sqrt{300}} = 1.69$$

$CI = \left(\bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} , \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

$C.I = (54 - 1.69 , 54 + 1.69)$

$C.I = (52.31 , 55.69)$

---

## Testing of Hypothesis.

In many circumstances, to arrive at decisions about the population on the basis of sampling information, we make assumptions (or guesses) about the population parameters involved. such as assumption (or statement) is called a statistical Hypothesis which may or may not be true.

The procedure which enable us to decide on the basis of sample results whether a hypothesis is true or not is called Test of Hypothesis (or) Test of significance

Eg: 1) The majority of men in the city are smokers

2) The teaching method in both the schools are effective

Testing of Hypothesis:

1) Null Hypothesis ($H_0$): A null Hypothesis is the hypothesis which asserts that there is no significant difference between the statistic and the population parameter. This is denoted by $H_0$

Null Hypothesis $H_0$ : $\boxed{\mu = \mu_0}$

2) **Alternative Hypothesis ($H_1$)**: Any hypothesis which controductions the Null Hypothesis. It is denoted by $H_1$

i, $H_1 : \mu \neq \mu_0$

ii, $H_1 : \mu > \mu_0$ iii, $H_1 : \mu < \mu_0$

3) **Level of Significance ($\alpha$)**:

The level of significance denoted by $\alpha$ is the confidence with which we reject or accept the Null Hypothesis $H_0$.

We take either 5% (0.05) or 1 : (0.01) level of significance.

4) **Test of Statistic**:

There are several tests of significance with z, t, F etc. First we have to select the right test depending on the nature of the information given in the problem.

5) **Decision**: We compute the value of the appropriate statistic and ascertain whether the computed value falls in acceptance or rejection region depending on the specified level of significance.

⇒ if T.V > C.V then N.H. $H_0$ is Accepted (or) unbiased.

otherwise N.H $H_0$ is rejected (or) biased.

**Type I Error**: If N.H $H_0$ is true but it rejected by Test Procedure then the error made is called Type-I error

**Type II Error**: If N.H $H_0$ is false but it is accepted by test, then error committed is called Type - II error

---

1) A sample of 64 students have a mean weight of 70 kgs, can this be regarded as a sample from a population with mean weight 56 kgs and S.D 25 kgs

Sol: $\bar{x} = 70$ kgs . $n = 64$, $\mu = 56$ kgs , $\sigma = 25$ kgs

Step 1 : N.H $H_0 : \mu = \mu_0$

Step 2 : A.H $H_1 : \mu \neq \mu_0$

Step 3 : LS ($\alpha$) : 99% (0.01) $Z_{\alpha/2} = 1.96$ (T.V)

Step 4 : T.S : $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{70-56}{\frac{25}{\sqrt{64}}} = 4.48$ (C.V)

Step 5 : Conclusion: T.V < C.V N.H is rejected (or) biased.

---

2) An Oceanographer wants to check whether the depth of the ocean in certain region is 57.4 fattoms as had previously recorded. What can be conclude @ 0.05 level of the signification. If reading taken at 40 random location in the given region yielded a mean of 59.1 fattoms with a $\sigma$ of 5.2 fattoms.

Sol: Given $n = 40$, $\sigma = 5.2$, $\bar{x} = 59.1$, $\mu = 57.4$

Step 1 : N.H $H_0 : \mu = \mu_0$

Step 2: A.H $H_1 : \mu \neq \mu_0$

Step 3: L.S ($\alpha$): 5% (0.05) $Z_{\alpha/2} = 1.96$ (T.V),

Step 4: T.S $z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{59.1 - 57.4}{\frac{5.2}{\sqrt{40}}} = 2.06$ (C.V)

Step 5: Conclusion: T.V = 1.96 C.V = 2.06

CV > TV

N.H is Biased (or) rejected

3) A sample of 900 members has a mean of 3.4 cm and $\sigma = 2.61$ cm is this has been taken from a large population of mean 3.25 cm and $\sigma = 2.61$ cm if the population is normal & its mean is unknown Find the 95% fiducial limits of true mean.

Sol: **Note:** Let we know the mean of the sample, population, deviation, size of the sample

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Fiducial limits (or) Confidence limits,

$$C.I = \left(\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$$

$n = 900$, $\bar{x} = 3.4$, $\mu = 3.25$, $\sigma = 2.61$

Step 1: N.H $H_0 : \mu = \mu_0$

Step 2: A.H $H_1 : \mu \neq \mu_0$

Step 3: L.S ($\alpha$): 5% (0.05) $Z_{\alpha/2} = 1.96$ (T.V)

Step 4: T.S $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{3.4 - 3.25}{2.61/\sqrt{900}} = 1.72$ (C.V)

Step 5: Conclusion:

TV > CV

N.H is accepted (or) unbiased.

Confidence interval $= \left(\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$

$= \left(3.4 - 1.96 \times \frac{2.61}{\sqrt{900}}, 3.4 + 1.96 \times \frac{2.61}{\sqrt{900}}\right)$

$= (3.22, 3.57)$

4) A sample of 400 items is taken from the population whose $\sigma = 10$, the mean of the sample = 40; test whether the sample has come with mean 38 also calculate 95% of confidence interval of population.

Sol: $n = 400$, $\sigma = 10$, $\bar{x} = 40$, $\mu = 38$

Step 1: N.H $H_0 : \mu = \mu_0$

Step 2: A.H $H_1 : \mu \neq \mu_0$

Step 3: L.S ($\alpha$): 5% (0.05) $Z_{\alpha/2} = 1.96$ (T.V)

Step 4: T.S $Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{40 - 38}{10/\sqrt{400}} = 4$ (C.V)

Step 5: Conclusion T.V = 1.96, C.V = 4

T.V < C.V

NH is biased (or) rejected

$$C.I = \left(\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right)$$

$$C.I = \left(40 - 1.96 \frac{10}{\sqrt{400}}, 40 + 1.96 \frac{10}{\sqrt{400}}\right)$$

$$C.I = (39.02, 40.98)$$

5) An ambulance service claims that it takes an avg. less than 10 min to reach its destination in emerg. calls. A sample of 36 calls as a mean of 11 min. and the variance of 16 min test the claim at 0.05 level of significance.

Sol: $n = 36$; $\sigma^2 = 16$, $\sigma = n$; $\bar{x} = 11$; $\mu = 10$

Step 1 N.H $H_0 : \mu = \mu_0$

Step 2 A.H $H_1 : \mu \neq \mu_0$

Step 3 L.S ($\alpha$): 5% (0.05) $z_{\alpha/2} = 1.96$ (T.V)

Step 4 T.S $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{11-10}{4/\sqrt{36}} = 1.5$ (C.V)

Step 5 Conclusion T.V = 1.96 C.V = 1.5

T.V > C.V

It is unbiased.

6) It is claimed that a random sample of 49 tyres has a mean life of 15200 km. The sample was drawn from a population whose mean is 15100 km & $\sigma$ = 1200 km. T.S at 0.05 level of significance.

Sol: $n = 49$, $\bar{x} = 15200$, $\mu = 15150$, $\sigma = 1200$, T.S = 0.05

Step 1 : N.H $H_0 : \mu = \mu_0$

Step 2 : A.H $H_1 : \mu \neq \mu_0$

Step 3 : L.S ($\alpha$) : 0.05% i.e $z_{\alpha/2} = 1.96$ (T.V)

Step 4 : T.S $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{15200 - 15150}{1200/\sqrt{49}} = 0.29$ (C.V)

Step 5 : Conclusion T.V > C.V N.H is accepted (or) unbiased