PROBABILITY
AND
STATISTICS
BSE: 3rd
SEMESTER

22/1/24

Introduction To STATISTICS:

The word statistics is derived from Latin word "status" which means group of numbers, figures and facts. The subject statistics deals with the ① collection of data, ② representation of the data, ③ analysis (cause and effect) and ④ final conclusions.

→ Statistics (Behaviour).

Uses of Statistical Information

(By Graph)

① To inform the general public.

② To explain the things that have happened. (Match Summary)

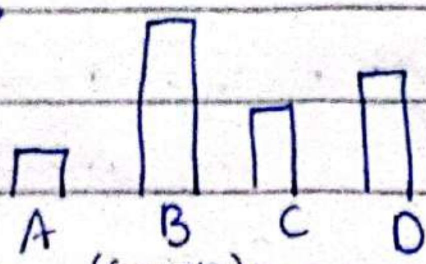
(Boys & girls marks)

③ To build general comparison.

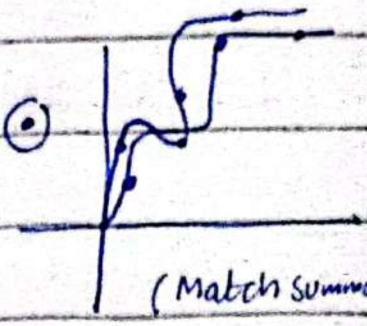
④ To predict the future outcome. (sale & advertisement).

⑤ To build association (finding relation b/w things)

Eg:



(Graph)



(Match summary)

- Marks Average comparison Boys & girls
- Sale = 3 + 10 adv
  = 3 + 10(5)
  sale = 53

(*) Integration:
  → Linear Algebra
  → Power Rule etc....

Variables and its Types:

Any characteristic which vary from person to person or individual to individual is known as variable. Eg: social status, CGPA, height, pocket money etc..... Opposite to variable is constant. (Semester, class, subject etc.....)

Types:

There are two types of variables:

1. Quantitative Variable
2. Qualitative Variable.

(a) Quantitative Variable:

Quantitative variables are numerically expressed.
Eg: Salary, Number of students or computers.

Quantitative variables is further divided into two types:
(i) Discrete (ii) Continuous.

(i) Discrete variables:

Discrete variables are obtained by counting. They cannot take decimal values in them. Eg: Number of chairs, Number of students, Number of classes. (Finding number of anything).

(ii) Continuous Variables:

Continuous variable is obtained by measurement. It can take infinite number of values between two points. E.g: Height, weight, CGPA, salary, marks etc.......

• Qualitative Variables:

Qualitative variables are expressed by labels and names. They are also known as attributes (quality). Eg: social status, marital status, gender, color (eye, hair etc----)

-> Scales and its types in Next class

-> Book Name:
Introduction to Statistical Theory By Chaudhary Sher Muhammad (Part I & II).

24/1/24

(*) Aggregate of all outcomes having same characteristics is known as **population**. E.g: All students of PUCIT, All families in Lahore, All Iphone Users.

population

characteristics

The numerical quantity obtained from population is known as **parameter**.

(*) The subset of population is known as **sample**. E.g: Blood in the test tube, 20% of students of PUCIT, 1/4th of the chocolate bar.

The numerical quantity obtained from sample is known as "**Statistics**"

(Relation) → Statistics is obtained from sample.

⊙ Types of Statistics:

There are two types of statistics:

1. Descriptive Statistics
2. Inferential statistics.

(*) In **descriptive** statistics, the goal is to describe. E.g:

(i) For respresentation of the data we use tabular and graphical method.

(ii) For representing the data by single value we use mean, median, mode, variance and standard deviations.

(iii) For representing the relationship we use "correlation".

(*) In **Inferential** statistics, we deal with the cause and effect, justification of the claim, future predictions and precautions regarding population on the basis of sample population. Eg:

(i) For showing the dependency and cause and effect we use regression.

(ii) For justification of the claim we use hypothesis testing.

(iii) For comparison we use ANOVA (Analysis of variance). or t-test.

(iv) For finding the unobserved variables (confidence) we use factor analysis.

Level of measurement:

In research, we measure the variable according to its category, magnitude and limits.

There are four types of scales:

1. Nominal scale
2. ordinal Scale
3. Interval Scale
4. Ratio Scale.

(*) Nominal Scale:

The variables which are categorized by names only are measured by nominal scale. E.g: Gender, Religion, CNIC, Roll No. etc.... Names

(*) Ordinal Scale:

The variables which have identity property and magnitude property are measured in ordinal scale. E.g: Social status (upper, middle, lower) Grades (A, B, C, D), outcome of a match (win, lose, tie).

(*) Intervals Scale:

The variables which have identity property, magnitude property and doesnot have minimum value are measured in interval scale.

: salary (zero not exist) , marks (negative not exists)

Eg: Temperature (Negative value exist) , IQ Score (PPSC test), SAT Score.

Ratio Scale:

The variables which have identity property, magnitude property and have minimum value of zero are measured in ratio scale.

E.g: Salary, Weight, Height etc....

29/1/24

* SAMPLING TECNIQUES:

The process of drawing sample from the population is known as sampling or sample experiment.

There are two types of sampling:

1. Random Sampling
2. Non-Random Sampling

RANDOM SAMPLING:

In random sampling every element of population has equal chance of occurrence. It is free from researcher biased process. It follows some rules and regulations:
There are four types of random sampling:

1. Simple Random Sampling
2. Stratified Random Sampling
3. Cluster Random Sampling
4. Systematic Random Sampling

Simple Random Sampling:

In simple random sampling, the population is homogeneous (having same characteristics). In this sampling process, the population unit are known. We select the sample by

: Jeeto Pakistan (homogeneous audience)
: Gold Game (Simple Random Sampling)

lottery method.

⊛ Stratified Random Sampling:
When the population is heterogeneous, first we make homogeneous group known as "stratra", then we select samples from each group. This process is known as stratified sampling. Eg: All students of PUCIT are not studying SE.

⊛ Cluster Random Sampling:
When the population is divided into different geographical areas, then we randomly select a cluster and draws sample from it. Eg: Election areas are cluster. Areas wise data collection.

⊛ Systematic Random Sampling:
When the population is arranged in column wise

or row wise situation then we select the sample in a series form. This is known as systematic sampling.
E.g: Anarkali shops arranged in rows and columns. Data collecting from different shops. Series wise data collection. (4, 8, 12, 16, 20, 24 ...)

⊙ NON-RANDOM SAMPLING:

When we select the sample by our convenience or personal biased process. This kind of sampling is known as non-random sampling.

⊙ MEASURE OF CENTRAL (origin) TENDENCY:

For representing the data by a single value we use central value (origin of the data). By central value we mean, "mean", "median", "mode".

⊙ MODE:

: trend (People talks much)

Mode is known as the trending value or most frequently occuring value. It is used or it is suitable for nominal scale. E.g: Name, Gender etc.....

For the rest of the data, it is used to find shape.

E.g: Blood Group of 10 students are given below:

A, B, O, O, B, A, B, B, O, A

The mode is "B" because it occurs four times.

⊙ MEDIAN:

Median is used for ordinal data. It represents

: Next class Assignment

∴ shape of data (where data is flowing)

the middle score of the data, also contributes in the shape of the data. **Steps of median:**

(i) Arrange the data

(ii) Note the number of observations, if the observations are odd then median will be equal to:

$$\text{ODD} \Rightarrow \text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{obs}$$

If the observation is even, then median will be equal to:

$$\text{EVEN} \Rightarrow \text{Median} = \frac{1}{2}\left[\left(\frac{n}{2}\right)^{th} \text{obs} + \left(\frac{n}{2}+1\right)^{th} \text{obs}\right]$$

31/1/24

(*) Median is also suitable for open ended data and for leaning data (Skewed).

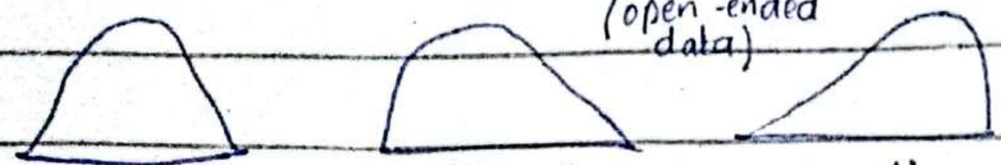
E.g: The following are the weights of the students and the data is leaning:

47, 41, 52, 43, 56, 35, 49, 55, 42

which measure of central tendency is suitable for data?

We use median because data is leaning.

(open-ended data)



Symmetry

Positive Skewed (+ skewed) (Starting value not known)

Negative skewed (- skewed) (ending value not known).

→ 35, 41, 42, 43, 47, 49, 52, 55, 56

$N = 9$

$$\text{Median} = \left(\frac{N+1}{2}\right)^{th} \text{obs}$$

$$= 5^{th} \text{ obs} = 47$$

⊛ For the following data compute median:

10, 19, 54, 80, 15, 16

① → 10, 15, 16, 19, 54, 80

② → $N = 6$

$$\text{Median} = \frac{1}{2}\left[\frac{N}{2}^{th} \text{obs} + \left(\frac{N}{2}+1\right)^{th} \text{obs}\right]$$

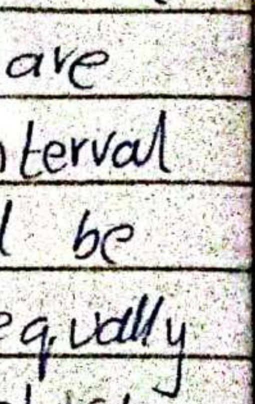
$$= \frac{1}{2}\left[3^{rd} \text{ obs} + 4^{th} \text{ obs}\right]$$

$$= \frac{1}{2}(16+19) = \frac{1}{2}(35) = 17.5$$

⊛ MEAN:

Mean is known as the balancing point. It is suitable for the variables which are measured in ratio and interval scale but they should be symmetry which means equally distributed. If the variables are qualitative, we first code them in numeric terms and then find the mean.

There are three types of mean:



1. Arithmetic mean
2. Geometric mean
3. Harmonic mean

⊛ Arithmetic Mean:

Arithmetic mean

is also known as average. It is widely used in practical and social studies.

Formula:

$$\bar{X} = \frac{\Sigma X}{N}$$: $\frac{\text{Sum}}{\text{No. of Observations}}$

② Geometric Mean:

Variables which are obtained by the result of rate or ratio. Eg: Speed, Velocity and acceleration etc.....

Formula:

$$G.M = \sqrt[N]{\pi X}$$: π (product of data)

When the data is in 100s, we can't find it.

$$G.M = \sqrt[N]{\pi X}$$

$$= (\pi X)^{\frac{1}{N}}$$

$$= \frac{1}{N} \log(\pi X)$$

$$= \frac{1}{N} \Sigma \log X$$

$$\boxed{G.M = \text{antilog}\left(\frac{1}{N}\Sigma \log x\right)}$$

⊛ Harmonic Mean:

Harmonic Mean is also used when the variable is obtained by the result of rate or ratio but the numerator will be fixed.
E.g: Finding the speed of 10 vehicles for 50 Km.
Formula:

$$\boxed{H.M = \frac{N}{\Sigma \frac{1}{x}}}$$

⊙ Practice Questions

① ⊛ Which measure of central tendency is suitable for the height of five students:

6.5, 6.1, 5.9, 5.8, 5.2

→ Arithmetic Mean is suitable for this data because the

: No information about leaning
: So mean will be consider.

variable is measured in ratio scale and it is not the resultant of rate or ratio.

$$A.M = \frac{(6.5+6.1+5.9+5.8+5.2)}{5}$$

$$\boxed{A.M = \frac{(29.5)}{5} = 5.9 \text{ Height}}$$

2. Which measure of central tendency is suitable for speed of 10 vehicles:

30, 35, 32, 40, 42, 45, 30, 32, 35, 32

$\rightarrow \log 30 + \log 35 + \log 32 + \log 40 + \log 42 + \log 45 + \log 30 + \log 32 + \log 35 + \log 32$

$= 15.43$

$$G.M = \text{antilog}\left(\frac{1}{N} \Sigma \log X\right)$$

$$= \text{antilog}\left(\frac{15.4638}{10}\right)$$

$$= \text{antilog}(1.54638)$$

$$\boxed{G.M = 35.19151}$$

3. Which measure of central tendency is suitable for

gravitational force of 5 objects from 10 m.

$\rightarrow$ 9.7, 9.9, 9.8, 10.1, 10.3

$N = 5$

$$\Sigma \frac{1}{X} = 9.7^{-1} + 9.9^{-1} + 9.8^{-1} + 10.1^{-1} + 10.3^{-1}$$

$$= \frac{1}{9.7} + \frac{1}{9.9} + \frac{1}{9.8} + \frac{1}{10.1} + \frac{1}{10.3}$$

$$= 0.5022$$

$$\boxed{H.M = \frac{5}{0.5022} = 9.956}$$

Relation b/w G.M, H.M, A.M:

$$H.M \leq G.M \leq A.M$$

Q:4 : Balancing Point (A.M)
Middle Score (G.M)
(Trend) (H.M)

Measure of Dispersion:

Measure of dispersion defines the spread

(mean, median, mode)

of data around central value.
We will only study variance and standard deviation because they are used in practical statistics.

① Variance:

Variance is defined as distance square (distance^2) around the mean. Whereas the **standard deviation** is the positive square root of variance.

Formula of Variance and standard deviation:

⊙ Variance $= \underset{(sigma)}{\sigma^2} = \frac{\Sigma(X-\bar{X})^2}{N}$: 8 steps

$= \sigma^2 = \frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2$: 3-4 steps

⊙ Standard Deviation $= \sigma = \sqrt{\frac{\Sigma(X-\bar{X})^2}{N}}$

$\sigma = \sqrt{\frac{\Sigma x^2}{N} - \left(\frac{\Sigma x}{N}\right)^2}$

AR : Book is must

For the following data compute the variation (variance) and the spread (standard deviation) around the mean :

9, 7, 11, 13, 2, 4, 5, 5

→ Variance(Variation) $= 6^2 = \frac{\Sigma(X-\bar{X})^2}{N}$

① $\bar{X} = \frac{\Sigma X}{N}$

② X

③ $(X - \bar{X})^2$

① Short method :

$$6^2 = \frac{\Sigma x^2}{N} - \left(\frac{\Sigma X}{N}\right)^2$$

$\Sigma X = 9+7+11+13+2+4+5+5$

$\Sigma x^2 = 9^2+7^2+11^2+13^2+2^2+4^2+5^2+5^2$

$\Sigma x^2 = 490$

$$6^2 = \frac{(490)}{8} - \left(\frac{56}{8}\right)^2$$

variance $= 6^2 = 12.25$

→ Standard Deviation $= \sqrt{6^2}$

$= \sqrt{12.25}$

$6 = 3.5$

12/2/24

* Coefficient of Variation:

For finding which dataset is more consistent or reliable, we use coefficient of variation. It is represented as C.V:

$$C.V = \frac{S.D}{Mean} \times 100$$

: SD = standard Deviation

The dataset which have the smallest C.V will be the most reliable and consistent. The dataset which have the largest mean will be the better one but C.V determines the consistency. (PCB and ICC example). **E.g:**

- N = 13 (Houstan)

$\Sigma x = 75+71+64+56+53+55+47+55+52+50+50+50+47$

$\Sigma x = 725$

Variance = 72.7928

- Pittsburgh:

$\Sigma x = 539$

Variance = 81.941

$\odot$ Mean $= \frac{\Sigma X}{N} = \frac{725}{13}$ (Houstan)

$= 55.77$

$\odot$ Mean $= \frac{\Sigma X}{N} = \frac{539}{13}$ (Pittsburgh)

$= 41.461$

Houstan mean is more but it is better while pittsburgh is consistent because of C.V.

: (Probability Topic)

* **Sample Space:**

The term **experiment** refer to the planned activity which leads towards result of Known outcomes. (Cooking, Writing results).

A random experiment is that experiment whose outcomes are not pre-determined. (Ludo).

A single performance of an experiment or random experiment is Known as trial.

A set consisting of all possible outcomes of a random experiment is known as sample space. E.g: Outcome of flipping a coin (Head or tail) (outcomes).

Sample Space of a Well Shuffle Deck (cards):

→ Total Cards = 52

Red = 26

Black = 26

Shapes = ♠ ; ♥ , ♦ , ♣

♠	♥	♦	♣
Spade (13)	Heart (13)	Diamond (13)	Club (13)

Face = King, Queen, Jack

King	Queen	Jack
4	4	4

→ Numbers:

2	3	4	5	6	7	8	9	10	
4	4	4	4	4	4	4	4	4	= 36

(Not a number) Ace = 4

Sample Space of Two Dice:

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

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

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

⊙ Sample Space of Two Coins and Three Coins:

Coin 1: H, T

$: 2^n$
$\rightarrow n:$ outcomes
$: 2^2 = 4$

Coin 2: H, T

$= \{ (HH), (HT), (TH), (TT) \}$

Coin 3: H, T

$: 2^3 = 8$

(combining it by up one (first write with head then tail).

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

Coin 4: H, T

$: 2^4 = 16$

$= \{ HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH$

THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT}

* Event:

Event is known as the subset sample space. E.g:

1. Even Numbers in a dice = {2, 4, 6}
2. Same Numbers on two dice = {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}
3. Exactly having two heads in three coins: = {HHT, HTH, THH}
4. Exactly three heads in four coins. = {HHHT, HTHH, THHH, HHTH}
5. Having 3 or 4 digit in a well shuffled deck: $\underset{(3)\ \text{card}}{4} + \underset{(4)\ \text{card}}{4} = 8$.

* Counting Rules of Sample Space: (selecting)

There are three basic counting rules for finding the sample space of large outcomes.

→ **Rule-1:**

Rule of Multiplication.

If a compound experiment consist of two experiments such that the first experiment has "m" outcomes and the second experiment has "n" outcomes, then the compound experiment will have "m×n" outcomes. (Dice and coin) (2x6 outcomes)

Eg: Outcome of a dice and coin will be $6 \times 2 = 12$.

→ **Rule-2:**

Rule of Permutation.

If we have "n" distinct objects and we want to select a subset of "r" with the restriction of order, then we use permutation method which is defined as:

$$^{n}P_r = \frac{n!}{(n-r)!}$$: Do it by calculator.

By order we mean position of the outcomes.

(Key Points) → In questions we identify permutation by the words of 'order', 'ranking', 'rating', 'research' (medical research) etc.... (other research behaviour) (we use combination)

→ **Rule-3:**

Rule of Combination.

If we have "n" distinct objects and we want to select a subset of "r" without the restriction of order, then we use combination method which is defined as: (No position representation)

$$^{n}C_r = \frac{n!}{r!(n-r)!}$$

Comparison b/w Permutation and Combination:

Suppose we have 4 outcomes

$= \{A, B, C, D\}$, we randomly want to select two outcomes from them in order and without order. List all possible outcomes of both cases.

→ In order: (Permutation)

A, B, C, D : 2 select outcomes

AB	BA	CA	DA
AC	BC	CB	DB
AD	BD	CD	DC

: $^4P_2 = 12$

→ Without order:

: No behaviour repeats by same person

AB	BC
AC	BD
AD	CD

: $^4C_2 = 6$

: AB = BA (so, no repetition)

*: → Presentation (5th week). (Integration).

14/2/24

EXAMPLES

① A particular cellphone company offers four models of phone, each in six different colors and each available with any one of fire calling plans. How many combinations are possible?

Sol:-

: Rule of Multiplicatio

→ Combinations $= 4 \times 6 \times 5$

$= 120$

② The foreign language club is showing a four movie marathon of subtitled movies. How many ways can they choose four from the eleven available.

Sol:-

Combinations $= {}^{11}C_4$: ways (so combination).

$= 330$.

③ How many different ways can a city health department inspector visits 5 restaurant in a city with 10 restaurant.

(Inspection)(So Permutation).

Sol:-

Permutation $= {}^{10}P_5$

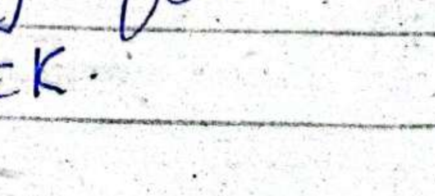
$= 30240$

4 There are two major roads from city X to city Y and four major roads from city Y to city Z. How many different trips can be made from city X to Z passing through city Y.

Sol:-

Trips $= 2 \times 4$: Rule of multiplication

$= 8$



5 Select randomly four cards from a deck.

Sol:-

Combination $= {}^{52}C_4$

$= 270,725$

(6) Selecting top 3 players from 11 members of baseball.

Sol:-

$$\text{Permutation} = {}^{11}P_3$$
$$= 990$$

(7) How many ways can a dinner pattern select 3 appetizers and 2 main course, if there are 6 appetizers and 5 main course on the menu.

Sol:-

$$A \Rightarrow 6 \Rightarrow 3$$
$$M \Rightarrow 5 \Rightarrow 2$$

$$\text{Ways} = {}^{6}C_3 \times {}^{5}C_2$$
$$= 20 \times 10$$
$$= 200$$

(8) The environment air protection agency investigates 9 mills. In how many

ways a representor select 5 from them to investigate in 1 week.

Sol:-

Permutation $= 9P_5$
$= 15120$

(9) In a board of director, having 8 peoples, how many ways we can have one chief officer, one director and one treasurer.

Sol:-

: Specification (order)

Permutation $= 8P_3$
$= 336$

(10) How many different signals can be made by using atleast 3 different flags, if there are 5 different flags, from which to select.

Sol:-

At least $3 = {}^5C_3 + {}^5C_4 + {}^5C_5$
$= 16$

: atleast maximum combination

* Probability:

Probability is Known as chance of occurence it defines the measure of uncertainity. It lies between 0 and 1. It could be represented as fraction, percentage, decimals and ratio form (:).

Probability of an event is defined as:

$$P(E) = \frac{n(E)}{n(S)}$$

: $\frac{n(E)}{n(S)}$

$$P(E) = \frac{\text{no. of outcomes in event}}{\text{no. of outcomes in sample space}}$$

: Dice (Same)
$= \frac{6}{36} = \frac{1}{6}$
$= 1.66$
$= 16.67\%$

→ Properties of Probability:

(i) Probability lies between 0 and 1. E.g: Probability of even number in a dice.

$P(even) = \frac{3}{6} = \frac{1}{2} = 0.5$

(ii) Probability of uncertain (non-existence) event will be zero.
E.g: Probability of having 7 in a dice.

(iii) The sum of all probabilities of a random experiment will be one.
E.g: ⊙ Dice:

	1	2	3	4	5	6	
(occurence chance)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$= 6/6 = 1$

⊙ Coin:

H	T	
$\frac{1}{2}$	$\frac{1}{2}$	$= 1$

(iv) Probability of complimentary event is defined as:
$P(\bar{A}) = 1 - P(A)$
E.g: The probability of not having 4 in a dice.
$P(\bar{4}) = 1 - P(4)$
$= 1 - \frac{1}{6} = \frac{5}{6}$

EXAMPLES

① Find the probabilities of following events in two dice:

(a) A sum of 9 (add)
(b) A sum of 7 or 11
(c) Same numbers (doubles)
(d) A sum less than 9
(e) A sum greater than or equal to 10.

② Selecting a card from a deck:

(a) A Queen
(b) A club
(c) A face card
(d) A 3 or 7
(e) A diamond or club
(f) An Ace.

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$\underbrace{\quad}_{\frac{1+2+3+4+5+6+5}{36}}$

Question : 1

Sample space of two dice

= { (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)
}

(a) Probability $= \frac{4}{36}$: sum of 9

(b) Probability $= \frac{6}{36} + \frac{2}{36}$

(c) Probability $= \frac{6}{36}$

(d) Probability $= \frac{26}{36}$

(e) Probability $= \frac{6}{36}$

QUESTION : 2

(a) Probability of Queen $= \frac{4}{52}$

(b) Probability of club $= \frac{13}{52}$

(c) Probability of face card $= \frac{12}{52}$

(d) Probability of 3 and 7 $= \frac{8}{52}$: This procedure for (ii) six card only

(e) Probability of diamond or club $= \frac{26}{52}$

(f) Probability of an Ace $= \frac{4}{52}$

(3) Randomly select two cards from a deck. Find the probabilities of following:

(a) Both are diamond

(b) Both are Queens

(C) Both are Red.

Total Sample Space $= {}^{52}C_2 = 1326$

(a) $P(\text{Both of diamonds}) = \dfrac{{}^{13}C_2}{{}^{52}C_2}$

$=$

(b) $P(\text{Both Queen}) = \dfrac{{}^{4}C_2}{{}^{52}C_2}$

$=$

(c) $P(\text{Both Red}) = \dfrac{{}^{26}C_2}{{}^{52}C_2}$

$=$

** Quiz Next week Wednesday (After assignment syllabus).

19/2/24

Example

• Probability Using Combination Method:

→ Find the probability of selecting two queens from a well-shuffled deck.

Sol:-

$$\text{Total} : {}^{52}C_2$$

$$P(2\text{ Queens}) = \frac{{}^{4}C_2}{{}^{52}C_2}$$: 4 Queens

: 52 cards

$$=$$

→ In a box there are 12 bulbs, 3 of which are defective, if 4 are selected at random. Find the probabilities of following:

(a) Zero defective bulb
(b) One defective bulb
(c) Three defective bulbs

Total : 12 defective = 3 : Nd = 9

select = $x = 4$

$$\text{Total Sample Space} = {}^{12}C_4$$

(a) $P(0 \text{ defective}) = \dfrac{{}^{9}C_4 \times {}^{3}C_0}{{}^{12}C_4}$

$$=$$

(b) P(1 defective): $\frac{{}^{3}C_1 \times {}^{9}C_3}{{}^{12}C_4}$

=

(c) P(3 defective) = $\frac{{}^{3}C_3 \times {}^{9}C_1}{{}^{12}C_4}$: 9 (Nd)

=

→ Three cards are randomly selected without (combination) replacement. Find the probability of following:

(a) ALL are blacks

(b) ALL are spade

(c) ALL are face cards

: without means combination

Total = ${}^{52}C_3$

(a) P(all Black): $\frac{{}^{26}C_3}{{}^{52}C_3}$

(b) P(all spade) = $\frac{{}^{13}C_3}{{}^{52}C_3}$

(c) P(all face cards) = $\frac{{}^{12}C_3}{{}^{52}C_3}$

** → Topic (Sample space, Event counting Rules and Probability) (Next Class Quiz)

*** Mutually Exclusive Event:**

Two or more events are said to be mutually exclusive if they cannot occur at the same time (We can say that they don't have any common terms). E.g: ① Having an even number and an odd number in a dice.

② Having an Engineering student and a medical student.

If the events have something common or they can occur at the same time then these events are known mutually exclusive. E.g: ① Having an even number and a multiple of 3 in a dice.

Even: 2, 4, 6] mutually Non-Exclusive
Multiple of 3: 3, 6]

→ Having 4th semester student and a statistics. (Non-Mutually).

* Addition Rule:

Addition rule depends on the concept of mutually exclusive and non-mutually exclusive to identify the rule we have the words like "either" and "or".

When two events are mutually exclusive then the probability of having either "a" or "b" will be defined as:

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

or ($\cup$) Union

When two events are

non-mutually exclusive, then the probability of either 'a' or 'b' will be defined as:

: And ($\cap$)

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

Examples

→ In Lahore, there are 4 Tim Hortons cafe, 5 Gloria Jeans and 6 coffee planets. Find the probability of having either "Tim Hortons" or "Gloria Jeans cafe" for having coffee.

: Law of Addition

Sol:-

$$P(\text{Tim H} \cup GJ) = P(\text{Tim H}) + P(GG)$$

$$= \frac{4}{15} + \frac{5}{15} = \frac{9}{15}$$

→ In a fish tank, there are 24 gold fish, two angel fish and 5 guppies. If a fish is randomly selected.

Find the probability that it is gold fish or angel fish.

Sol:-

$P(\text{Goldfish} \cup \text{Angel fish})$

$$= P(\text{Gold fish}) + P(\text{Angel fish})$$

$$= \frac{24}{31} + \frac{2}{31} = \frac{26}{31}$$

→ A card is randomly selected from a deck. Find the probabilities of following:

(a) A four <u>or</u> a diamond

(b) A club or a diamond

(c) A Jack or a black

(a) A four or a diamond:

A 2 3 4 5 6 7 8 9 10 K Q J

♡

◇

♠

♣

$$\rightarrow P(4 \cup D) = P(4) + P(D) - P(4 \cap D)$$

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

(b) A club or a diamond:

$$\frac{13}{52} + \frac{13}{52} = \frac{26}{52}$$

(c) A Jack or a black:

$$P(\text{Jack} \cup \text{Black}) = P(\text{Jack}) + P(\text{Black}) - P(B \cap J)$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52}$$

$$= \frac{28}{52}$$

→ In a book shelf, there are 100 adults books and 160 children books in which seventy of adult books are non-fiction while sixty of children books are non-fiction. If a book is randomly selected. Find the probabilities of following:

(a) Fiction books

(b) Not a children non-fiction book

(c) An adult book or a children non-fiction book.

Sol:-

	Adults	Children	
Non-Fiction	70	60	130
Fiction	30	100	130
			260

(a) $P(\text{Fiction}) = \frac{130}{260} = \frac{1}{2}$

(b) $P(\underline{\text{Not}} \text{ children Non-fiction}) = 1 - P(\text{children Non Fiction Books})$ ∵ complimentary property

$= 1 - \frac{60}{260}$

$= \frac{200}{260}$

(c) $P(\text{Adult book or Children non-fiction book}) = P(\text{adult}) + P(C.NF)$

$= \frac{100}{260} + \frac{60}{260} = \frac{160}{260}$

$= \frac{8}{13}$

→ In a statistics class, there are 18 seniors and 12 juniors. 6 of seniors are female and 8 of the juniors are male. Find the probabilities of following:

(a) A junior or a female
(b) A senior or a male
(c) A junior or a senior.

	Female	Male	
Junior	4	8	12
Senior	6	12	18
	10	20	30

$P(J) + P(F) - P(J \cap F)$

(a) $P(\text{Junior or Female}) = \frac{12}{30} + \frac{10}{30} - \frac{4}{30}$

$= \frac{18}{30}$

(b) $P(\text{senior or male}) = \frac{18}{30} + \frac{20}{30} - \frac{12}{30}$

$= \frac{26}{30}$

(c) $P(\text{Senior or a junior}) = \frac{18}{30} + \frac{12}{30} - \frac{0}{30}$

$= \frac{30}{30} = 1$

21/2/24

(*) In combination (without order) BA = AB while in permutation (order) BA ≠ AB.

(*) Arrangement (Counting Rules)

* Independent and Dependent Event:

When two or more events doesnot effect the chance of occurence of other events or each other. Then these events are independent events. E.g: Tossing a coin and selecting a card.

When the occurence of a event is affected by occurence of other events then these events are dependant events.

Eg: Selecting a girl than a boy from a class.
(For understanding the dependency in probability, the sample space is reduced).

Eg: In a class there are 35 students out of which 15 are girls and 20 are boys. Selecting two students at random without (defines dependency). replacement, find the probability that first one is girl and second one is boy.

Girls = 15

Boys = 20

Total = 35

$$P(G \cap B) = P(G) \times P(B/G)$$

$$= \frac{15}{35} \times \frac{20}{34}$$

- **Multiplication Rule:**

When two events

are independent, the probability of both occuring is define as:

$$P(A \cap B) = P(A) \times P(B)$$: Both events are independent (with replacement).

E.g: Find the probability of having a head and a 6 in a dice.
(multiplication)

$$P(\text{Head} \cap 6) = P(\text{Head}) \times P(6)$$
$$= \frac{1}{2} \times \frac{1}{6} = \frac{1}{2}$$

→ Find the probability of having an even number in a dice and a face card from a deck. not a combination (we directly take outcomes)

$$P(\text{Even} \cap \text{Face}) = P(\text{Even}) \times P(\text{Face})$$
$$= \frac{3}{6} \times \frac{12}{52}$$
$$= \frac{36}{312}$$

→ When two events are dependent, then the chance of occurence is defined as:

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

such that / A is already occured

→ Find the probability of having 3 Jacks without replacement from a deck:

$$P(J_1 \cap J_2 \cap J_3) = P(J_1) \times P(J_2/J_1) \times P(J_3/J_1 \cap J_2)$$

$$= \frac{4}{52} \times \frac{3}{51} \times \frac{2}{50}$$

$$= \frac{1}{5525}$$

$$\frac{{}^{4}C_3}{{}^{52}C_3} = \frac{1}{5525}$$

***Condition Probability:** The probability that event B occurs such that event A is already occur is define as:

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

A and B could be the

common terms.

→ E.g: In a statistics class, there are 18 seniors and 12 juniors. 6 of seniors are female and 8 of the juniors are male.

	M	F	
S	12	6	18
J	8	4	12
	20	10	30

Find the probability of having a female such that she is senior:- (defines condition)

$$P(F/S) = \frac{P(F \cap S)}{P(S)}$$

$$= \frac{6/30}{18/30} = \frac{6}{18} = \frac{1}{3}$$

: such that after term always divide
: given that
: that
(words for conditional probability)

26/2/24

(*) "A" word then divide outcomes by total outcomes. Probability of

3 cards, 4 cards etc.... (combination)

* ⊛ Question 6.27 part (b and c) (Pg # 221) (law of addition)

-> 6.29 (part c) (law of addition).

-> 6.30 (part b and c).

-> 6.33 (part b), 6.34 (part a, c)

-> 6.35 (part a, b).

-> 6.43 (part a, b), 6.36 (part c)

EXAMPLES

-> The gift basket store had the following pattern of gift baskets, contains the following combinations in the stock.

	Cookies	Mugs	Candy	
Coffee	20	13	10	43
Tea	12	10	12	34
	32	23	22	77

Find the probability of following:

(a) Tea such that having candy.

(b) Coffee and mug.
(c) Coffee such that it is in mug.
(d) Cookie or tea.

ⓐ $P(\text{Tea} | \text{Candy}) = \frac{P(\text{Tea} \cap \text{Candy})}{P(\text{candy})}$ (and)

$= \frac{12/77}{22/77} = \frac{12}{22}$

$= \frac{6}{11}$

ⓑ $P(\text{Coffee} \cap \text{Mug}) = \frac{13}{77}$

ⓒ $P(\text{coffee} | \text{Mug}) = \frac{P(\text{Coffee} \cap \text{Mug})}{P(\text{Mug})}$

$= \frac{13/77}{23/77}$

$= 13/23$

ⓓ $P(\text{Cookie} \cup \text{Tea}) = \frac{32}{77} + \frac{34}{77} - \frac{12}{77}$

$= \frac{32+34-12}{77}$

$= \frac{54}{77}$

Q (6.36)(part c)

→ Probability of having A = 0.60

$P(A) = 0.60$

$P(B) = 0.40$

$P(A \cap B) = 0.24$

$P(A/B) = ?$

$P(B/A) = ?$

$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{0.24}{0.40}$

$=$

$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{0.24}{0.60}$

$=$

→ Q 6:34 (part c)

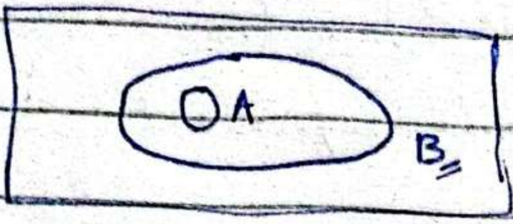
In a firm 20% of employees having accounting background while 5% of the employees are executives and have accounting background. If an employee has an accounting background what is the probability that employee is executive.

OP(Executive/Accounting)

→ $A \subset B$. (6·3) Theorem

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

: pg#191, 192



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

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

$B = A + (B - \bar{A})$

(6·4 Theorem

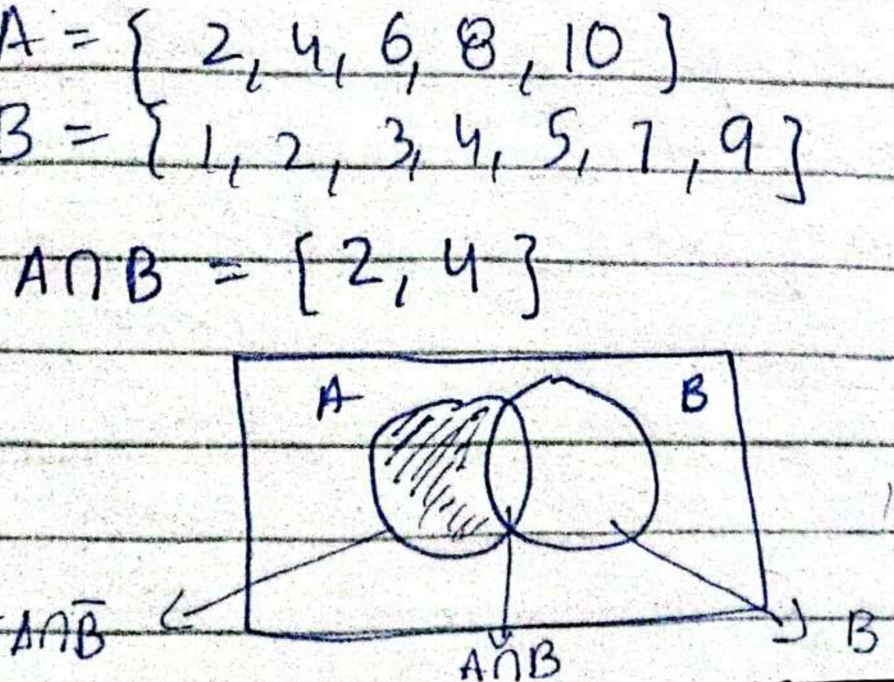
→ $P(A \cap \bar{B}) = P(A) - P(A \cap B)$

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

$A = \{2, 4, 6, 8, 10\}$

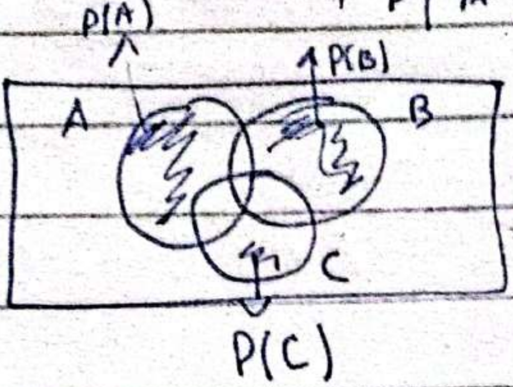
$B = \{1, 2, 3, 4, 5, 7, 9\}$

$A \cap B = \{2, 4\}$



→ Theorem 6.6 (pg # 195) :

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



Eg: A, B, C are the following sets:

$A = \{2, 3, 4\}$ $S = \{2, 3, 4, 5, 6, 7\}$

$B = \{3, 4, 5\}$

$C = \{5, 6, 7\}$

Find the probability of having either A or B or C.

$P(A) = 3/6$ $P(B \cap C) = \frac{1}{6}$

$P(B) = 3/6$ $P(A \cap B \cap C) = 0$

$P(C) = 3/6$

$P(A \cap B) = 2/6$

$P(A \cap C) = 0/6 = 0$

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

$$= \tfrac{3}{6} + \tfrac{3}{6} + \tfrac{3}{6} - \tfrac{2}{6} - 0 - \tfrac{1}{6} - 0$$

$$= 1$$

→ $(A+B+C)^2 = A^2 + B^2 + C^2 + 2AB + 2BC + 2CA$

$(a+b+c)(a+b+c) = a^2 + b^2 + c^2 + 2ac + 2bc + 2c$

→ $(A-B-C)(A-B-C)$

⊛ Atleast and Utmost:

(If word come then add)

(Minimum to Maximum).

(Maximum

→ Atleast = Minimum Limit to the Maximum point.

→ Utmost = Minimum to the limit.

→ 6.21 (Part a) (Pg #220).

Total eggs = n = 12

Bad eggs = 2, Good = 10

Randomly $= x = 4$: randomly (combination) : If "a" then

① $P(1 \text{ is bad}) = \frac{{}^{2}C_1 \; {}^{10}C_3}{{}^{12}C_4}$

② At least 1 is Bad $= P(1 \text{ is Bad}) + P(2 \text{ is Bad})$

$= \frac{{}^{2}C_1 \times {}^{10}C_3}{{}^{12}C_4} + \frac{{}^{2}C_2 \times {}^{10}C_2}{{}^{12}C_4}$

→ **6.24** :

Total balls $= n = 14$

Red $= 4$

Black $= 5$

White $= 5$

Randomly $= x = 6$: $14 - 4 = 10$

① $P(3 \text{ are red}) = \frac{{}^{4}C_3 \times {}^{10}C_3}{{}^{14}C_6}$

② At least 2 white $=$ 2 white + 3 white + 4 white + 5 white

$= \frac{{}^{5}C_2 \times {}^{9}C_4}{{}^{14}C_6} + \frac{{}^{5}C_3 \times {}^{9}C_3}{{}^{14}C_6} + \frac{{}^{5}C_4 \times {}^{9}C_2}{{}^{14}C_6} + \frac{{}^{5}C_5 \times {}^{9}C_1}{{}^{14}C_6}$

→ **6.25** (Part b) :

Men $= 6$, Women $= 8$

$r = 5$

$n = 14$

Sol:-

Men = $0, 1, 2, \boxed{3, 4, 5}$

Women = $5, 4, 3, \boxed{2, 1, 0}$

→ At least 3 Men = 3 Men + 4 Men + 5 Men.

→ 6.26 :

Total: ${}^{10}C_3$

(a) $\frac{{}^{4}C_3}{{}^{10}C_3}$

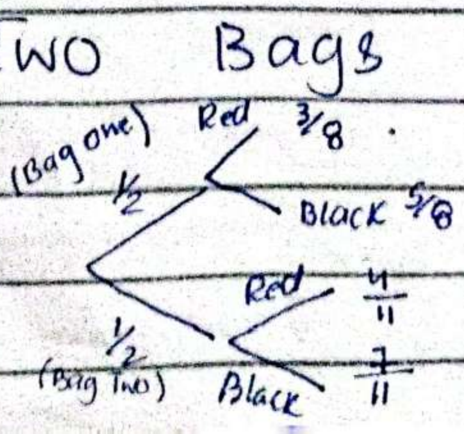
(b) $\frac{{}^{6}C_3}{{}^{10}C_3}$

(c) At least = 1 + 2 + 3

→ 6.43 (a) :

: Part b (Make Tree of it)!

Two Bags



5|3|24

(*) Book page #224 (6.50, 6.51).

→ **Q # 6.50:**

$$P(A) = 1/6 \qquad P(\bar{A}) = 5/6$$

$$P(B) = \frac{1}{4} \qquad P(\bar{B}) = 3/4$$

$$P(C) = \frac{1}{3} \qquad P(\bar{C}) = 2/3$$

$$\therefore P(X=1) = P(A \cap \bar{B} \cap \bar{C}) + P(\bar{A} \cap B \cap \bar{C}) + P(\bar{A} \cap \bar{B} \cap C)$$

$$= \left[\frac{1}{6} \times \frac{3}{4} \times \frac{2}{3}\right] + \left[\frac{5}{6} \times \frac{1}{4} \times \frac{2}{3}\right] + \left[\frac{5}{6} \times 3/4 \times \frac{1}{3}\right]$$

$$= 0.397$$

$$\therefore P(\text{first hit}) = P(A \cap \bar{B} \cap \bar{C})$$

$$= \left[\frac{1}{6} \times \frac{3}{4} \times \frac{2}{3}\right] = 0.083$$

→6.51:

$P(A) = 0.4 \quad P(\bar{A}) = 0.6$

$P(B) = 0.5 \quad P(\bar{B}) = 0.5$

$P(C) = 0.6 \quad P(\bar{C}) = 0.4$

(a) $P(\text{all hit}) = P(A \cap B \cap C)$

$= 0.4 \times 0.5 \times 0.6$

$= 0.12$

(b) ~~At least 1~~

$P(X=1) + P(X=2) + P(X=3)$

→ $P(X=1) = P(A \cap \bar{B} \cap \bar{C}) + P(\bar{A} \cap B \cap \bar{C}) + P(\bar{A} \cap \bar{B} \cap C)$

$= (0.4 \times 0.5 \times 0.4) + (0.6 \times 0.5 \times 0.4) + (0.6 \times 0.5 \times 0.6)$

$= 0.38$

→ $P(X=2) = P(A \cap B \cap \bar{C}) + P(\bar{A} \cap B \cap C) + P(A \cap \bar{B} \cap C$

$= (0.4 \times 0.5 \times 0.4) + (0.6 \times 0.5 \times 0.6) + (0.4 \times 0.5 \times 0.6)$

$= 0.38$

→ $P(X=3) = P(A \cap B \cap C)$

$= (0.4 \times 0.5 \times 0.6) = 0.12$

$\rightarrow P(X=1) + P(X=2) + P(X=3) = 0.88$

: Short way

$\rightarrow 1 - P(\text{all not hit})$

$1 - P(\bar{A} \cap \bar{B} \cap \bar{C})$

$1 - (0.6 \times 0.5 \times 0.4)$

$1 - 0.12$

0.88

(iii) One hits target.

$$P(X=1) = P(A \cap \bar{B} \cap \bar{C}) + P(\bar{A} \cap B \cap \bar{C}) + P(\bar{A} \cap \bar{B} \cap C)$$

$$= (0.4 \times 0.5 \times 0.4) + (0.6 \times 0.5 \times 0.4) + (0.6 \times 0.5 \times 0.6)$$

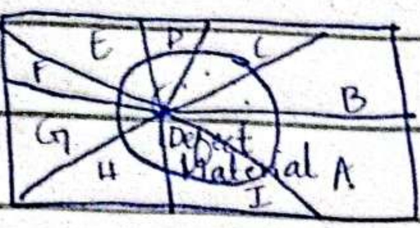
$$= 0.38$$

(iv)

$\rightarrow$ Bayes Theorem

It is also known as inverse propability theorem. When a sample space is further divided into sub-sample space. It is quite difficult to calculate the probabilities of sub-sub events.

E.g: Different brands of clothing have defective material in their lot. Finding the probability of defective will be quite difficult because first we will select brand and find defect in that brand.



$$\rightarrow P(A_i/B) = \frac{P(A_i)\,P(B/A_i)}{\Sigma P(A_i)\,P(B/A_i)}$$

→ **Eg:** In a factory, machine A, B, C manufacture 25%, 35% and 40% of the total output respectively of their

: pg #213 : 6.61

outputs 5%, 4% and 2% respectively are the defective material. If a material is selected at random. What is the probability that (Defective) the material comes from machine A, machine B and machine C.

→ Sol:-

A $P(A) = 0.25 \rightarrow P(D/A) = 0.05$

B $P(B) = 0.35 \rightarrow P(D/B) = 0.04$

C $P(C) = 0.4 \rightarrow P(D/C) = 0.02$

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

$$= (0.25 \times 0.05) + (0.35 \times 0.04) + (0.4 \times 0.02)$$

$$= 0.0345$$

$$\rightarrow P(A/D) = \frac{P(A) \times P(D/A)}{\text{Total Defective}} = \frac{0.25 \times 0.05}{0.0345}$$

$$= 0.3623$$

$$\rightarrow P(B/D) = \frac{P(B) \times P(D/B)}{\text{Total Defective}} = \frac{0.35 \times 0.04}{0.0345}$$

$$=$$

$\rightarrow P(C/D) = \dfrac{P(C) \times P(D/C)}{\text{Total Defective}}$

$=$

$\rightarrow$ **6.61:**

$P(A) = 0.02$ $50\% = 0.5$

$P(B) = 0.03$ $30\% = 0.3$

$P(C) = 0.05$ $20\% = 0.2$

Probability Distribution:

A tabular arrangement of outcomes of random experiment along with their corresponding probabilities is known as probability distribution.

E.g: Outcomes of a dice:

X = outcomes of dice	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

The sum of probabilities is always equal to 1.

→ The mean of probability distribution is known as expected value of X.

$$E(X) = \Sigma x \cdot P(X=x)$$

→ Variance of outcome is defined as:

$$Var(X) = E(X^2) - [E(X)]^2.$$

: 7.2 (part b)
: Example 7.1 & 7.2

$$= \Sigma x^2 P(X=x) - [\Sigma x P(X=x)]^2$$

$\rightarrow S.D_x = \sqrt{Var(x)}$

(Imp for Exam) **E.g:** Find the mean and variance of having head in three coins.

X = head in 3 coins

X	$P(X=x)$	$x P(X=x)$	$x^2 P(X=x)$
0	1/8	$0 \times 1/8 = 0$	$0 \times 1/8 = 0$
1	3/8	$1 \times 3/8 = 3/8$	$1^2 \times 3/8 = 3/8$
2	3/8	$2 \times 3/8 = 6/8$	$4 \times 3/8 = 12/8$
3	1/8	$3 \times 1/8 = 3/8$	$9 \times 1/8 = 9/8$
	1	12/8	24/8

HHH
HHT
HTH
HTT
THH
THT
TTH
TTT

$$E(x) = \Sigma x P(X=x) = \frac{12}{8} = \frac{3}{2} = 1.5$$

$$Var(x) = \Sigma x^2 P(X=x) - [\Sigma x P(X=x)]^2$$

$$= \frac{24}{8} - \left(\frac{12}{8}\right)^2$$

$$= 3 - 1.5^2$$

$$= 0.75$$

$$S.D_x = \sqrt{0.75}$$

$$\boxed{S.D_x = 0.8660}$$

18/3/24

(*) Mean $\Rightarrow E(X) = \Sigma x P(X=x) = 12/8$

Variance$(X) = E(X^2) - [E(X)]^2 =$

$= \Sigma x^2 P(X=x) - [\Sigma x P(X=x)]^2$

$SD(X) = \sqrt{\Sigma x^2 P(X=x) - [\Sigma x P(X=x)]^2}$

(*) Linear Combination of Random Variable:

By linear combination, we mean adding or subtracting any constant to the variable. There are the following properties of mean, variance and standard deviation:

(1) If we add or subtract any constant to the variable, same constant will be added or subtracted in the arithmetic mean i.e: $Y = a + X$ then

$\bar{Y} = a + \bar{X}$

Eg: suppose the following observations are the outcomes of 5 weights: 45, 50, 55, 60, 75.

Make a coded variable by subtracting 5 from each observation, also compute mean of original and coded variable.

$X = 45, 50, 55, 60, 75$

$Y = X - 5$

$Y = 40, 45, 50, 55, 70$

$\rightarrow \bar{X} = \frac{\Sigma X}{N}$ | $\bar{Y} = \frac{\Sigma Y}{N}$

$= \frac{45+50+55+60+75}{5}$ | $= \frac{40+45+50+55+70}{5}$

$= 57$ | $= 52$

∴ Same effect of constant on mean.

② If we multiply and divide any constant to the observation same constant will be multiplied and divide in the arithmetic mean.

i.e: $Y = bX$ then $\bar{Y} = b\bar{X}$

If $Y = a + bX$ then

$\bar{Y} = a + b\bar{X}$

③ If we add or subtract any constant to the observation, there will be no effect on variance and standard deviation i.e:

$$Y = a + X \text{ then } Var\, Y = Var\, X$$
$$S.D\, Y = S.D\, X$$

Eg: Following are the 6 students marks of quiz:

$X: 10, 15, 12, 18, 20, 8$

suppose we add to each observation find the variance of both dataset.

$Y = 20, 25, 22, 28, 30, 18.$

$\rightarrow \Sigma X = 10 + 15 + 12 + 18 + 20 + 8 = 83$

$\Sigma X^2 = 10^2 + 15^2 + 12^2 + 18^2 + 20^2 + 8^2 = 1257$

$$\sigma^2 = \frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2 = \left(\frac{1257}{6}\right) - \left(\frac{83}{6}\right)^2$$
$$= 18.1389$$

$\rightarrow \Sigma Y = 20 + 25 + 22 + 28 + 30 + 18 = 143$

$\Sigma Y^2 = 20^2 + 25^2 + 22^2 + 28^2 + 30^2 + 18^2 = 3517$

: Expected (Average

$$= \left(\frac{3517}{6}\right) - \left(\frac{143}{6}\right)^2$$

$$= 18.1389$$

: No effect will be on SD

④ If we multiply or and divide any constant to the original variable standard deviation will be same times of that constant and variance will be constant square times. i-e:

~~$Y = bX$~~ ~~$\text{Var } Y = b^2 \text{ Var } X$~~

~~$Y = bX$~~ $S.D\, Y = b\, S.D\, X$ (Example provided in file)

If $Y = a + bX$ then $\text{Var } Y = b^2 \text{ Var } X$

$S.D\, Y = b\, SD\, X$

~~Q: If X and Y are independent random variables with variance~~ $\sigma^2 x = 5$ and ~~$\sigma^2 y = 3$~~, find the variance of the random variable $Z = -2X + 4Y - 3$

Sol:-

$$\boxed{Z = -2X + 4Y - 3}$$

$Var\ X = 5$

$Var\ Y = 3$

$$Var\ Z = 4\ Var\ X + 16\ Var\ Y$$
$$= (4 \times 5) + (16 * 3)$$
$$= 20 + 48$$
$$= 68$$

Q7:

X	0	1	2	3
$P(X=x)$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
$X\,P(X=x)$	0	$\frac{1}{8}$	$\frac{2}{2}$	$\frac{3}{8}$

$$g(X) = 2X + 3$$
$$E(g(x)) = 2E(X) + 3$$
$$= 2 \times \frac{12}{8} + 3$$
$$=$$

Q: 4.36 :-

0	0.4
1	0.3
2	0.2
3	0.1

Mean and

Variance.

Q: $y = (x-1)^2$

$y = x^2 - 2x + 1$

$E(y) = E(x^2) - 2E(x) + 1$

20|3|24

E.g ④ A box contains 4 red, five green and six yellow balls. Four balls are randomly selected. Make the probability distribution of having green balls.

Sol:-

Red : 4

Green : 5

Yellow : 6

Total : 15

Randomly = 4

X: Green balls	$P(X=x)$
0	$\frac{{}^5C_0 \times {}^{10}C_4}{{}^{15}C_4} = 0.15$
1	$\frac{{}^5C_1 \times {}^{10}C_3}{{}^{15}C_4} = 0.43$
2	$\frac{{}^5C_2 \times {}^{10}C_2}{{}^{15}C_4}$
3	$\frac{{}^5C_3 \times {}^{10}C_1}{{}^{15}C_4}$
4	$\frac{{}^5C_4 \times {}^{10}C_0}{{}^{15}C_4}$

* Discrete Probability Distribution:

A variable which is measured by counting can't take infinite number of values between two points is known discrete variable. The probability distribution of discrete variable is known as discrete probability distribution.

* Binomial Distribution:

When the outcome of random experiment consists only two complimentary events. E.g: Head and tail, success and failure, right and wrong, good and defective etc....

The probability of having an event is define by "p": and it's complement is defined by "q" : $\boxed{q = 1 - p}$

The experiment is run for independent fixed trials. Fixed trials are denoted by "n".

The mean of binomial distribution is "np". The variance of binomial distribution is "npq" and the standard deviation of binomial distribution is "$\sqrt{npq}$".

Formula for finding the probabilities in binomial distribution is:

$$\boxed{P(X=x) = {}^{n}C_{x}\, p^{x} q^{n-x}}$$

E.g: Grade $A = 40\%$

$\bar{A} = 60\%$

$n = 6$

$X = 3$

$\frac{A}{0.4}$ $\frac{}{0.6}$ $\frac{A}{0.4}$ $\frac{}{0.6}$ $\frac{A}{0.4}$ $\frac{}{0.6}$ (At any place A will come, so combination)

(0.6 3 times so q^{n-x}) (0.4 3 times so p^{x} or)

Eg. 40% of Americans think that college degree is not important for business decisions. 5 persons are randomly selected. Find the probabilities of following:

(a) Exactly 2 will agree with the statement.

(b) Utmost 2 will agree with the statement.

(c) Atleast 4 will agree with the statement.

(d) Fewer than 2 will agree with the statement.

(e) Greater than 4 will agree with the statement.

(f) Find the mean, variance and standard deviation of the statement.

Sol:-

$P = 0.4 \quad q = 0.6 \quad n = 5$

(a) $P(X=2) = {}^nC_x \, p^x \, q^{n-x}$

$= {}^5C_2 (0.4)^2 (0.6)^{5-2}$

$= 5C_2 \, (0.4)^2 (0.6)^3$

$P(X=2) = 0.3456$

* : Chap # 8
(Example (8.1, 8.2, 8.3, 8.4, 8.5, 8.8, 8.10, 8.7)

ⓑ $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2)$

atmost

$= {}^5C_0\, 0.4^0 0.6^5 + {}^5C_1\, 0.4^1 0.6^4 + {}^5C_2\, 0.4^2 0.6^3$

$=$

ⓒ $P(X \geq 4) = P(X=4) + P(X=5)$

$= {}^5C_4\, 0.4^4 0.6^1 + {}^5C_5\, 0.4^5 0.6^0$

$=$

ⓓ $P(X < 2) = P(X=0) + P(X=1)$

$=$

$=$

ⓔ $P(X > 4) = P(X=5)$

$=$

$=$

ⓕ Mean $= np = 5 \times 0.4$

Variance $= npq = 5 \times 0.4 \times 0.6$

$S.D = \sqrt{npq} = \sqrt{5 \times 0.4 \times 0.6}$

$=$

Eg:

* If the 60% of the voters in a Larger district prefer candidate A. What is the probability that in a sample of 12 voters exactly

7 will be present A.

⊛ Poisson Distribution:
(chances less)

It is an extension of binomial distribution. It is applied when chances of having an event is less than $\boxed{5\%}$ or the sample size is greater than $\boxed{20}$. Its mean and variance are equal to
Mean $= x - p = \mu = \lambda$(mean) , $\sigma^2 = \mu = np$

The formula for finding the probability in poisson distribution is :

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

$$\sigma^2 = \mu = np.$$

Eg: Chances of having a wrong print is one in 500. Find the probability that exactly 4 prints will be wrong in 10000 pages.

Sol:-

$$P = \frac{1}{500} = 0.002 , n = 10000$$

$$X = 4$$

Less than 0.05 so poisson is applied.

$\mu = np = 10000 \times 0.002 = 20$

$P(X=4) = \frac{20^4 \cdot e^{-20}}{4!}$

$= 0.0000137$

Eg: The manufacture of pins knows that there will be 3% defective material in manufacturing a pins. If he sent the lot of 200 pins. Find the probability that utmost 3 will be defective one.

Sol:-

$p = 3\% = 0.03$

$n = 200$

$\mu = np = 200 \times 0.03 = 6$

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

$= \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} + \frac{6^2 e^{-6}}{2!} + \frac{6^3 e^{-6}}{3!}$

$=$

21/3/24

*Normal Distribution:

It is continuous variable distribution in which Mean, Median, Mode are equal.

→ Mean, Median, Mode are equal.

→ It ranges from +∞ to -∞ where we plot a graph of normal variable, there is no gap between points. The slope of the graph is bell shaped. Its peak point is the mode.

→ The graph cannot touch the x-axis because of its range

→ The shape of the bell is symmetrical (right and left side equal).

→ We cannot find a single point probability in normal distribution because the variable is of continuous variable.

→ Formula for the finding probability of normal distribution is defined as:

$$P(X<x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

→ We cannot use this formula because it will take 4-5 pages instead of this we will use standard normal distribution table. For this we transform x to z score.

$$z\text{ Score} = \frac{x-\mu}{\sigma}$$

⊙ STEPS SNDT :

: Standard Normal Distribution Table

① Identify mean, variance and standard deviation from the statement.

② Rewrite the probability in symbols.

③ Transform x to z Score.

④ See the table:

(i) The table contains the values of less than probabilities (Cumulative probabilities).

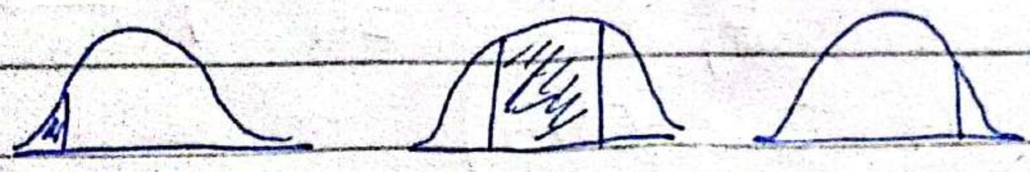
(a) If the probabilities is of less than, use direct table. $P(z<x)$: use direct table.

(b) If the probabilities is of greater than :

$$P(z>x) = 1 - P(z<x).$$

(c) If the probabilities is in between two points then:

$$P(a<z<b) = P(z<b) - P(z<a)$$



25/3/23

★ EXAMPLE:

A ~~service~~ survey states that: "Women spend an average of 1460 dollars on cosmetic with the standard deviation 24. If the variable is normally distributed. Find the probabilities of following:

(a) Women spend less than 143 dollars.

(b) Women spend greater than 155 dollars.

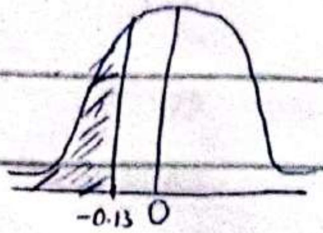
(c) Women spend in between 140 and 155 dollars.

Sol:-

$\mu = 146 \quad \sigma = 24$

(a) $P(X < 143)$

$P\left(\underset{\text{(zScore)}}{\frac{X-\mu}{\sigma}} < \frac{143-146}{24}\right)$



$P(Z < -0.13)$

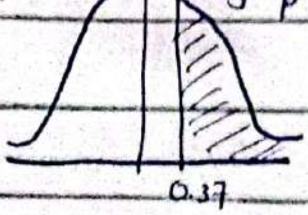
$= 0.4483$

: More to table (0.1 → 0.5517)

(b) $P(x > 155)$

(Draw area under graph)

$P\left(\frac{X-\mu}{\sigma} > \frac{155-146}{24}\right)$



$P(Z > 0.37)$

$1 - P(Z < 0.37)$

$1 - 0.6443$

$= 0.3557$

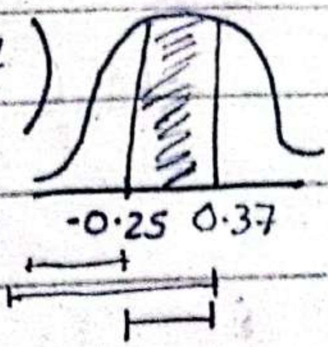
(2nd Rules of table).

pg # 390
9.23

(c) $P(140 < x < 155)$

$P\left(\frac{140-146}{24} < z < \frac{155-146}{24}\right)$

$P(-0.25 < z < 0.37)$



$P(z < 0.37) - P(z < -0.25)$

$0.6443 - 0.4013$

0.243

: symbols carefully use $(<, >)$

9.23:

Sol:-

$\mu = 500 \quad \sigma = 100$

(a) $P(X > 650)$

(b) $P(X < 250)$

(c) $P(325 < X < 675)$

(a) $P\left(\frac{X-\mu}{\sigma} > \frac{650-500}{100}\right)$

$P(z > 1.5)$

$1 - P(z < 1.5)$

$1 - 0.9332$

$= 0.0668$

(b) $P(X < 250)$

$$P\left(\frac{X-\mu}{\sigma} < \frac{250-500}{100}\right)$$

$$P(z < -2.5)$$

$$= 0.0062$$

③ Inverse Use of Table:

When we are asked to find the limit about certain point and its area is already given, we use inverse table to find the limit. **E.g:** If X is normally distribution with the mean 146 and variance 16. At which limit 95% of the data will be covered on the left side of the area.

Sol:-

$\mu = 146$; $\sigma^2 = 16$; $\sigma = 4$

: left (negative)
: right (positive)

$P(X < a) = 95\%$

$$P\left(z < \frac{a-146}{4}\right) = 0.95$$

Search for inner values on table

(Inverse not included)

$$\frac{a-146}{4} = 1.65$$

$$a = 146 + (1.65 \times 4).$$

Q: If the mean is 200 and standard deviation is 15. Find out the values of following:

(a) 25% data is below the value.
(b) 60% data is above the value
Imp ** (c) 75% data lies between two values in the middle.

Sol:-

$\mu = 200 \qquad \sigma = 15$

(a) $P(X < a) = 25\%$

$$P\left(z < \frac{a-200}{15}\right) = 0.25$$

$$\frac{a-200}{15} = -0.6745$$

$$a = (-0.6745 \times 15) + 200$$

$$\boxed{a = 189.88}$$

(b) $P(X > a) = 60\%$

$$P\left(z > \frac{a-200}{15}\right) = 0.60$$

$$1 - P\left(z < \frac{a-200}{15}\right) = 0.60$$

$$1 - 0.60 = P\left(z < \frac{a-200}{15}\right)$$

$$0.40 = P\left(z < \frac{a-200}{15}\right)$$

: 9.35, 9.31, 9.30, 9.33 (a, b), 9.35 (b).

$$-0.2530 = \frac{a-200}{15}$$

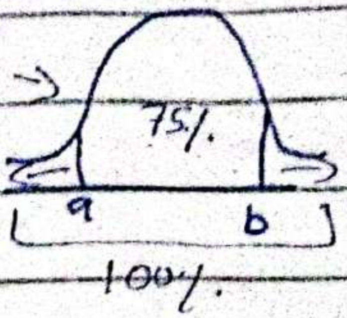
$$a = (-0.2530 \times 15) + 200$$

$$a =$$

100% = 1

ⓒ $P(a < X < b) = 75\%$.

$$P(X < a) + P(a < X < b) + P(X > b) = 1 \rightarrow$$



$$P(X < a) + 0.75 + P(X > b) = 1$$

$$P(X < a) + P(X > b) = 0.25$$

$\therefore P(X < a) = 0.125$; $P(X > b) = 0.125$

$P\left(Z < \frac{a-200}{15}\right) = 0.125$ | $1 - 0.125$ (Table) $P(X < b) = 0.875$

| $P\left(Z < \frac{b-200}{15}\right) = 0.875$

$\frac{a-200}{15} = -1.1503$ | $\frac{b-200}{15} = 1.1503$

$a = (-1.1503 \times 15) + 200$ | $b = (1.1503 \times 15) + 200$

⊙ 9.35 (b):

Find the mean and standard deviation if 31% of data ~~is~~ are under 45 and 8% data are over 64.

(interlinear corrections: "less than" above "are under"; "is above" above "are over")

→ Sol:-

∴ 80 Marks
(10 marks MCQs
25 marks S/QS (5 S/Q)
1 Long (Measure of Dispersion)
1 Long (Normal, Poisson, Binomial)
Ch#6 (S/Q)

$P(X<45)=0.31$

$P\left(z<\frac{45-\mu}{\sigma}\right)=0.31$

$\frac{45-\mu}{\sigma}=-0.4958$

$45=-0.4958\sigma+\mu \rightarrow (1)$

$P(X>64)=0.08$

$1-P(X<64)=0.08$

$P(X<64)=0.92$

$P\left(z<\frac{64-\mu}{\sigma}\right)=0.92$

$\frac{64-\mu}{\sigma}=1.4051$

$64=1.4051\sigma+\mu \rightarrow (2)$

$45=-0.4958\sigma+\mu$

$\pm 64=\pm 1.4051\sigma+\mu$

$-19=-1.9009\sigma$

$\sigma=\frac{19}{1.9009}$

$\sigma=9.9952 \rightarrow (3)$

put in (1) or (2) and get μ