Final

1. Exp mean, median & mode with eg:

→ Mean:

The mean is the sum of all values divided by number of values

Formula:

$$\text{Mean} = \frac{\text{Sum of all values}}{\text{No. of values}}$$

Eg: 10, 20, 30, 40, 50

$$\text{mean} = \frac{10+20+30+40+50}{5}$$

→ Mean = 30

→ Median:

The median is a middle value in a sorted list.
If even number of values, its avg of two middle values.

Eg: 10, 20, 30, 40, 50

⇒ Median = 30

Eg: 10, 20, 30, 40

⇒ $\text{Median} = \frac{20+30}{2} = 25$

→ Mode:-

The mode is the value that appears most frequently.

Eg: 10, 20, 20, 30, 40

⇒ Mode = 20

Use case:- Useful for categorical data or when finding most popular choice.

2 What is diff bet variance & standard deviation?

→ Variance	S.D
→ It shows avg square deviation from mean	→ It shows root of variance
→ Mathematical dispersion measure	→ Real world measure of spread.
→ Formula: $\frac{\sum (x_i - \bar{x})^2}{n}$ x_i : each value $\bar{x}$: mean of values n: total no. of value	→ Formula: $\sqrt{\text{Variance}}$

Eg: 10, 20, 30, 40, 50	Eg: 10, 20, 30, 40, 50
Variance = 250	S.D = 15.811
→ It shows how far each value in the dataset is from mean, on avg.	→ It shows square root of variance. It shows how much, on average, the value deviate from mean.

3 Define normal distrib & give one use case

→ A gaussian distribution or normal distribution is a bell shaped curve that shows how data is distributed

Most of the data points are concentrated around the mean, with fewer values farther away.

→ Practical use:
- → Height of people
- → Marks of student
- → Blood Pressure
- → IQ score

→ Mean, median & mode are equal
→ It is controlled by S.D.

→ Formula:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- μ = mean
- σ = standard deviation
- e = Euler number

Exp skewness & kurtosis

→ Skewness: It is statistical measure that describes asymmetry of a data distribution around its mean. It helps identify if data is symmetrically distributed or tilted

→ +ve skew :- Tail on right is long
-ve skew :- Tail on left is long
Symmetric :- Data is distributed

Formula :-

$$\frac{3(\text{Mean} - \text{Median})}{\text{Standard dev.}}$$

→ Kurtosis:

It measures the "tailedness" of a distribution - how heavy or light the tails are compared to a normal distribution. It tells us how likely it is to get outliers.

Types:
- → Leptokurtic
- → Mesokurtic
- → Platykurtic

→ If value > 3 - Leptokurtic
value < 3 - Platykurtic
value = 3 - Normal

Formula :-

$$\frac{E[(X-\mu)^4]}{\sigma^4}$$

→ μ - mean
σ - S.D.
E - expected value

5

5 What is Probability? Diff bet^n emplirical & theoretical Prob?

⇒ Probability is a measure of the likelihood that a particular event will occur.

→ It quantifies uncertainity using a number between 0 & 1.

→ 0 means the event can't happen.
1 means the event is certain

Formula:-

$$P(E) = \frac{\text{No. of favorable outcomes}}{\text{Total no. of Possible outcomes}}$$

→ P(E) is Probability of event
Favorable outcomes is outcome we want
Total outcome = all possible outcomes

→ Basic terms:-

Probability
Experiment
outcome
Sample space
Event
Favorable outcome

Theoretical	Emplirical
→ Theoretical probability is based on mathematic logic & assumes all outcomes are equally likely.	→ Emplirical Probability is based on real-world data or actual observations from experiments or datasets.
→ Formula: $P(E) = \frac{\text{No. of outcomes}}{\text{Total outcomes}}$	→ Formula: $P(E) = \frac{\text{No. of times event}}{\text{Total no. of trials}}$

6 Explain. independent & defendent event with 1 eg?

→ Independent:-

Two events A & B are independent if occurrence of one does not affect probability of other.

Formula:

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

Eg: Tossing coin & rolling die:

→ P(Head) = 0.5
P(4) = 1/6
P(Head & 4) = 0.5 × 1/6 = 1/12

→ Dependent:

Two events A & B are dependent if occurrence of one affects the probability of other.

Formula:

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

OR

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

Eg: Drawing cards without replacement.

→ First card is Ace: $P(A) = 4/52$
Second card is Ace: $P(B|A) = 3/51$
$P(\text{Two ace}) = 4/52 \times 3/51$.

7 What is intuition of baye's theorm

→ Baye's theorm is a way to update probabilities when you have new evidence.

It tells you how likely is given that another event has occured

Formula: $P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$

→ $P(A|B)$ - Probability of A given B happened
$P(B|A)$ - Probability of B given A happened
$P(A)$ - Probability of A
$P(B)$ - Probability of B

→ Intuition:

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

Prior: What you believed before
Likelihood: How probable the new evidence is
Evidence: Overall probability

8 Exp eigenvalue & eigenvector is simple term

→ & For a square matrix A:

$$Av = \lambda v$$

→ v = Eigenvector
λ = Eigenvalue

→ Multiplying a matrix by some vector does not change its direction only scales it by λ

→ Formula to find value:

$$(A - \lambda I)\, v = 0$$

=> Eigenvectors - Eigervectors are vectors that maintain their direction under a transformation.

=> Eigenvalue - Eigenvalues are values that scales vectors along that those direction.

→ In geometry:

stretching / compression along certain axes

→ In data science:

used in PCA (Principal component analysis)
→ eigenvectors = direction of max variance, eigenvalues = amout of var.