

Probability and Statistics

Introduction to Statistics

(decision making Process)

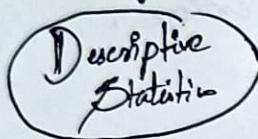
Def: Statistics is the science of collecting, organizing and analyzing data.

Data: "facts or pieces of information"

e.g. Height of students in classroom.

IQ of students in classroom.

Types of Statistics



It consists of organizing and summarizing data.

(1) Measure of Central tendency.

- (i) Mean
- (ii) Median
- (iii) Mode

(2) Measure of dispersion

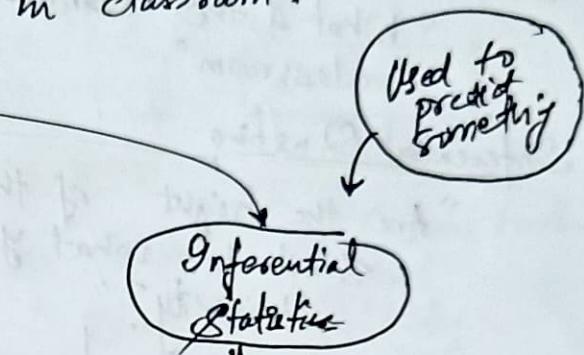
- (i) Variance
- (ii) Standard deviation

(3) Diff type of distribution of data.

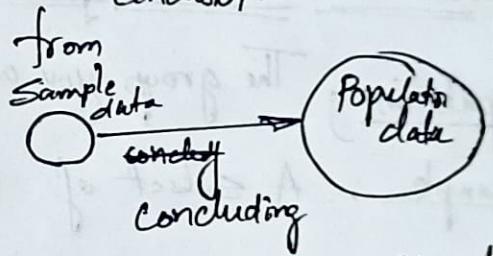
e.g. Histogram

Probability distribution function.

Probability Mass function.



It consists of using data you have measured to form conclusion.



- (i) Z-test
- (ii) t-test

Hypothesis Testing
 H_0 , H_1

- (iii) Chi square test

P value
Significance value

* for creating sample, be careful

- Sample size
- Random (No bias)
- Enough representative.

q: There are 20 Statistics classes at your University and you have collected the heights of student in the class. Heights are recorded as { 175, 180, 140, 140, 125, 160, 135, 190 }

Descriptive Question

"What is the ~~common~~ ^{avg} height of the entire classroom"

Inferential Question

"Are the height of the students in the classroom similar to what you expect in the entire University?"

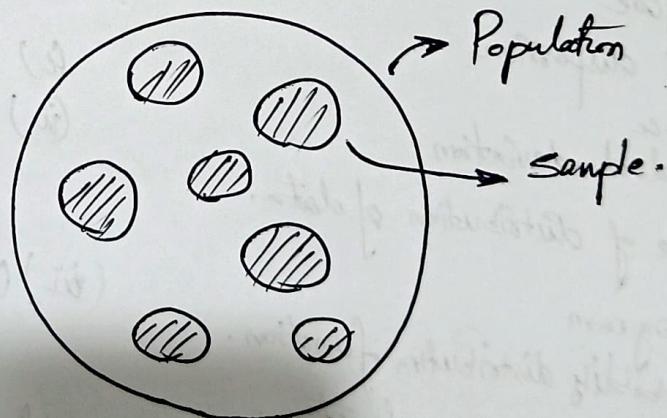
↓
population data

Population data & Sample data.

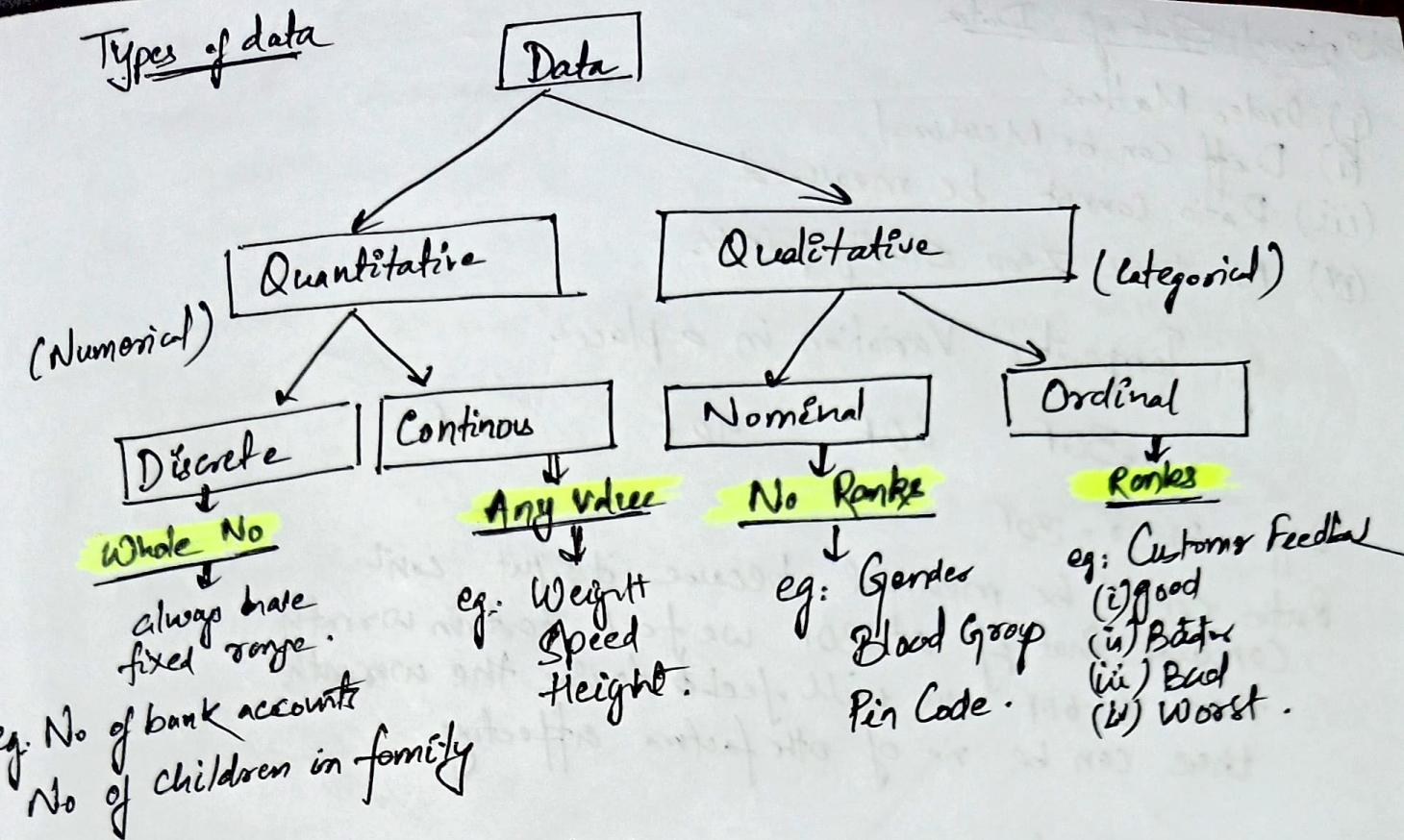
Population: The group you are interested in studying.

Sample: A subset of Population.

q: Exit Poll.



Types of data



Scale of Measurement of Data

- (1) Nominal Scale Data
- (2) Ordinal Scale Data
- (3) Interval Scale Data
- (4) Ratio Scale Data

(1) Nominal Scale Data

- (i) Qualitative
- (ii) Categorical Data
- (iii) Order/Rank doesn't matter

e.g.: Survey on favorite color.
(Here order does not matter on what color people like.)

(2) Ordinal Scale Data:

- (i) Ranking is important
- (ii) Order matters
- (iii) Difference cannot be measured

Just on the basis of rank we cannot calculate difference.

e.g.: An feedback form
(i) Best
(ii) Good
(iii) Bad
(iv) Worst

like in this case we don't know the reason for this ranking

(3) Interval Scale of Data

- (i) Order Matters
- (ii) Diff can be Measured
- (iii) Ratio cannot be measured
- (iv) No true Zero starting point.

e.g. Temperature Variation in a place.

30°F 60°F 90°F 120°F

$$\text{Diff: } 60 - 30 = 30^{\circ}\text{F}$$

Ratio cannot be measured because it's just count
conclude that if at 30°F we feel certain warmth
then at 60°F we will feel double the warmth
there can be no of other factors affecting.

(4) Ratio Scale Data

- ① The Order matters
- ② Diff are measurable (including ratios)
- ③ Common '0' starting point.

e.g. grades of students

0, 90, 60, 30, 75, 40, 50

Measure of Central Tendency

- (i) Mean or Average
- (ii) Median
- (iii) Mode

Mean

Population (N)

Primmer Mean

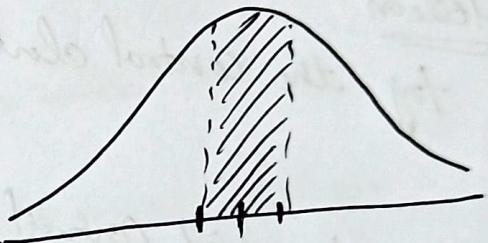
These is removal of certain values from bottom end and top to remove the impact of outliers.

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

$$\text{Population mean } (\mu) = \frac{\sum_{i=1}^N x_i}{N} = \frac{[1+1+2+2+3+3+4+5+5+6]}{10} = 3.2$$

Sample mean (n)

$$\text{Sample Mean } (\bar{x}) = \frac{\sum_{i=1}^n x_i}{n}$$



Median

$$X = \{4, 5, 3, 2, 1\}$$

Step

① Sort the random variable X . : $\{1, 2, 3, 4, 5\}$

② No of elements count : 5

③ if (count % 2 == 0)
find center element

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

or if (Count % 2 != 0)

find center element
 $\frac{2+3}{2} \rightarrow 3$ Median

$$\frac{2+3}{2} = 2.5 \text{ Median}$$

Why Median?

$$X = \{1, 2, 3, 4, 5\}$$

$$\bar{x} = \frac{1+2+3+4+5}{5} = 3$$

Shift / if now introduce another elem.

$$X = \{1, 2, 3, 4, 5, 100\}$$

(outlier)

because it does not belong to the distribution

$$\text{(mean)} \quad \bar{x} = \frac{1+2+3+4+5+100}{6} = \frac{115}{6} = 19.1$$

Now Median

for the central elem in ① Col $\rightarrow 3$

$$\textcircled{2} \text{ C.M.} = \frac{3+4}{2} = 3.5$$

* Median to find central tendency when outliers present.

Mode

frequency \rightarrow Max^m frequency.

$$\{2, 1, 1, 1, 4, 5, 7, 8, 9, 9, 10\}$$

Max^m frequency of an elem: 1 80, Mode = 1

Application of Mean, Median, Mode in Feature Engineering

Age	Weight	Salary	Gender	Degree
24	70	40K	M	BE
25	80	70K	F	-
27	95	10K	M	-
24	-	50K	-	PhD
32	-	60K	-	Masters
-	60	-	-	Bsc
-	65	55K	M	-
40	72	-	F	-

Numerical Value → We can find mean and fill null values.

↳ of outliers → median value.

Categorical Values: We can fill it with mode values.

Measure of Dispersion → (It describes the spread or variability of dataset.)
(Spread of the data)

- (1) Variance → (proportional to spread of the data tells us how much the values are far from spread.)
- (2) Standard deviation.

Variance

Population Variance (σ^2)

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

x_i = Data points

μ = population mean

N = population size

Sample Variance

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

↓
imp

x_i

\bar{x} → Sample mean

n = Sample size

Q Why we divide sample variance by $(n-1)$?

In order to create unbiased estimator of Bessel's corrⁿ

Q What is unbiased and biased estimator?

Unbiased: like expected value is equal to the true value of the parameter

Biased: expected value is not equal to the true value due to bias

e.g. If we pick 20 random students from a school and find avg

Unbiased } there is a good chance our estimator will give avg negt true to the actual height

Biased } If we pick up student only from the basketball team where students are typically taller our avg value will not be equal to true value.

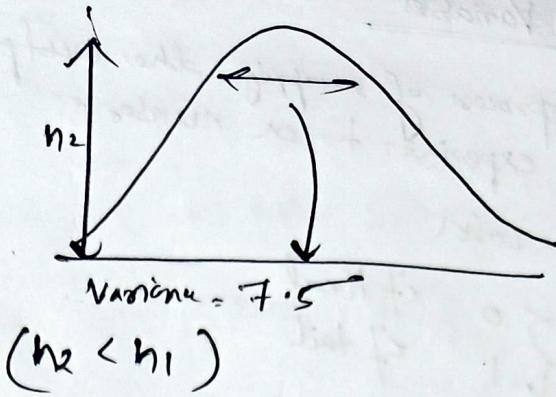
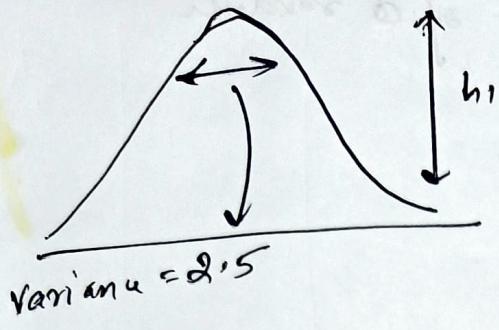
In Variance formula, we could have used mod, then square it, but we don't because inference conclusion is difficult for a population.

MAD (Mean Absolute Deviation)

$$\boxed{MAD = \frac{\sum_{i=1}^N |x_i - \bar{x}|}{n}}$$

less ~~more~~ prone + outliers like Variance still can't use to find result for a population.

Dispersion or Spread (Variance)



Standard Deviation

Population S.D

$$\sigma = \sqrt{\text{variance}} \quad (\sigma^2)$$

e.g. $\{x = \{1, 2, 3, 4, 5\}\}$

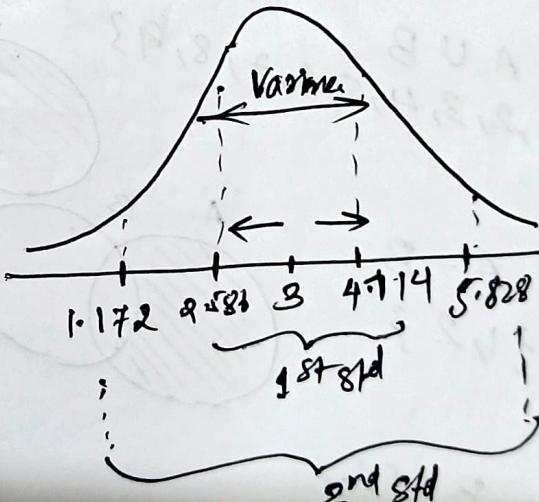
$$\bar{x} = 3$$

$$\sigma = 1.414 \quad \cancel{2.23}$$

Sample S.D

$$\text{Std} = \sqrt{s^2}$$

Sample variance



Coeff of Variation

$$CV = \frac{\text{Stand deviation} \times 100}{\text{mean}}$$

If it is basically volatility in a dataset, a measure of variability, how column which do not have same units

(High CV, high Variability)
(Low CV, low variability)

Random Variables

If it is the process of mapping the output of a random process or experiment to a number.

e.g. (1) Tossing a coin

$$x \begin{cases} 0 & \text{if Head} \\ 1 & \text{if tail} \end{cases}$$

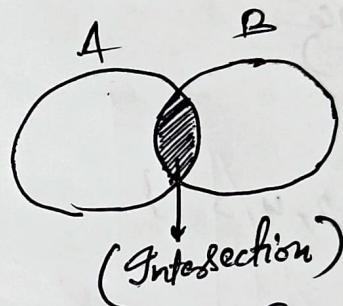
(2) Rolling a dice

Sets

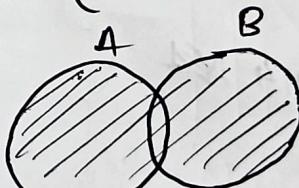
$$A = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

$$B = \{3, 4, 5, 6, 7, 8, 9\}$$

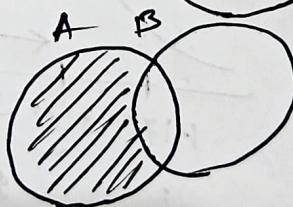
① Intersection: $A \cap B$
 $\{3, 4, 5, 6, 7, 8\}$



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



③ Difference: $A - B$
 $\{1, 2\}$



④ Subset:

$$\begin{aligned} A \rightarrow B &\rightarrow \text{False} \\ B \rightarrow A &\rightarrow \text{False} \end{aligned}$$

Histogram & Skewness → [Frequency]

Ages = {10, 12, 14, 18, 24, 26, 30, 35, 36, 37, 40, 41, 42, 43, 50, 51}

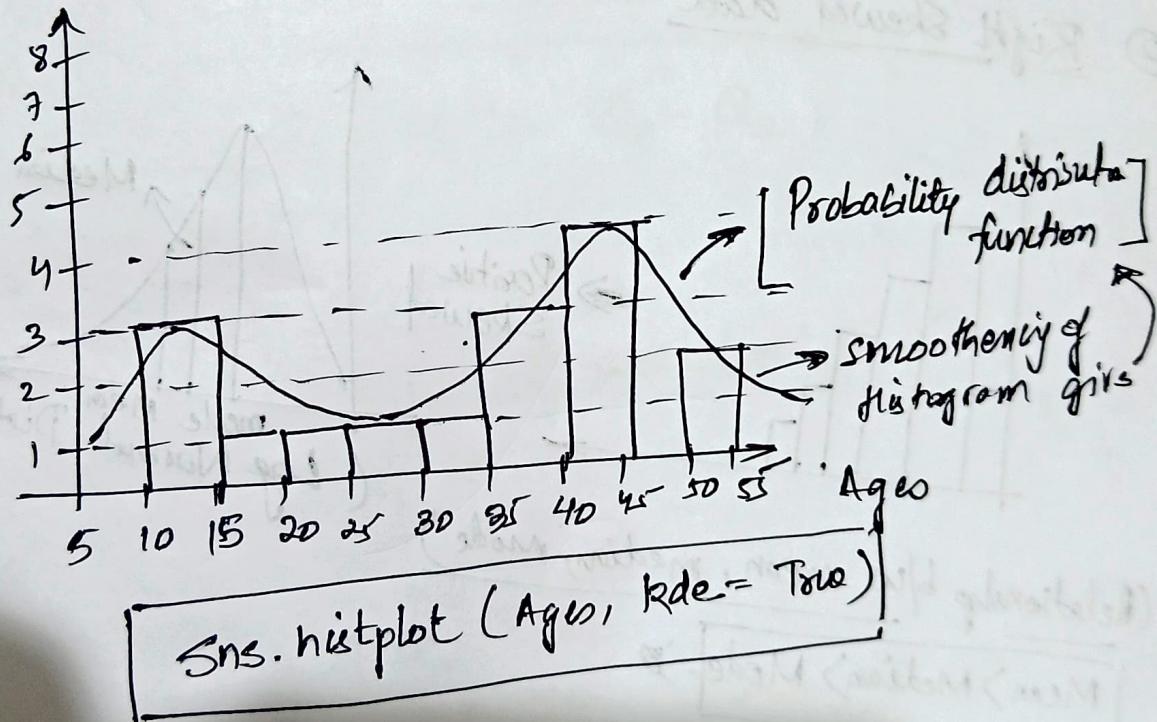
Consider, we have 60 values and we want 6 bins of 10 bins.

$$\frac{60}{10} = 6 \rightarrow \text{bin size} \quad (\text{No of bins} = 10)$$

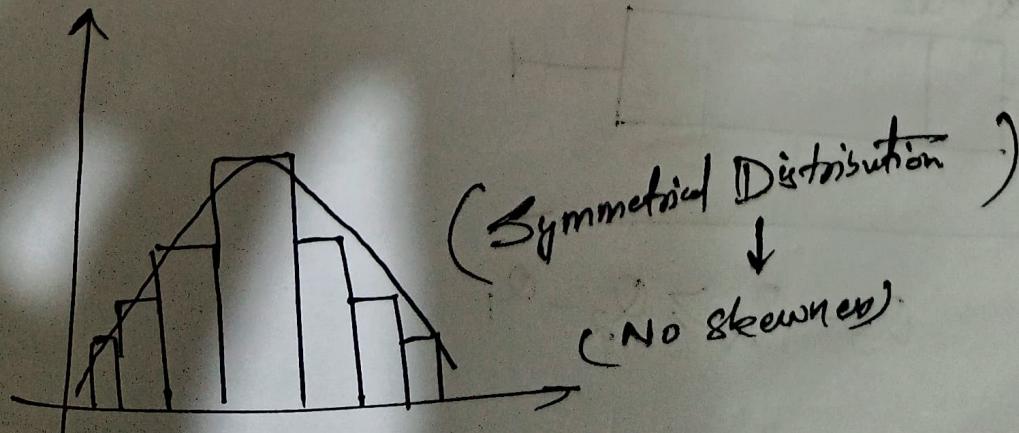
$$\frac{60}{20} = 3 \text{ bin size} \quad (\text{No of bins} = 20)$$

Histogram

Count.

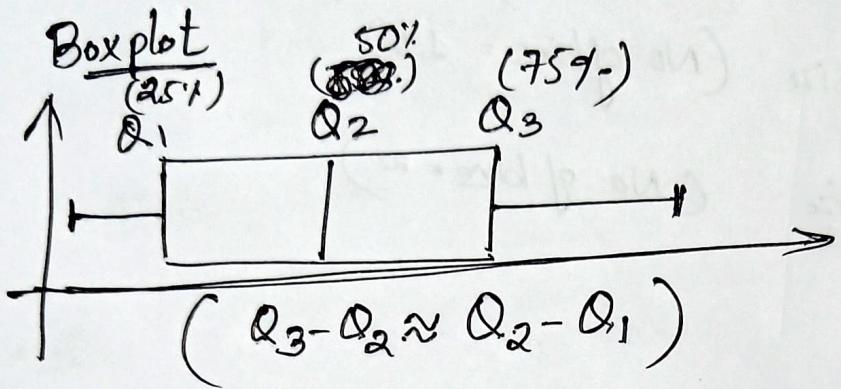


Symmetrical

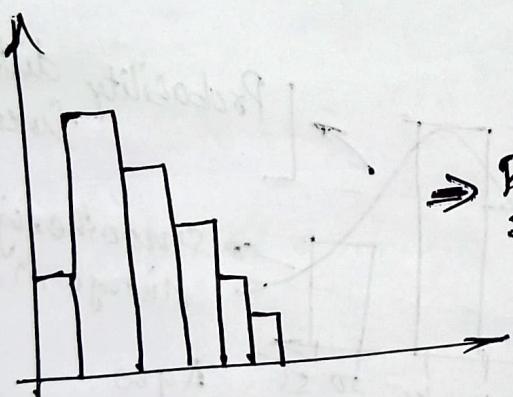


No skewed

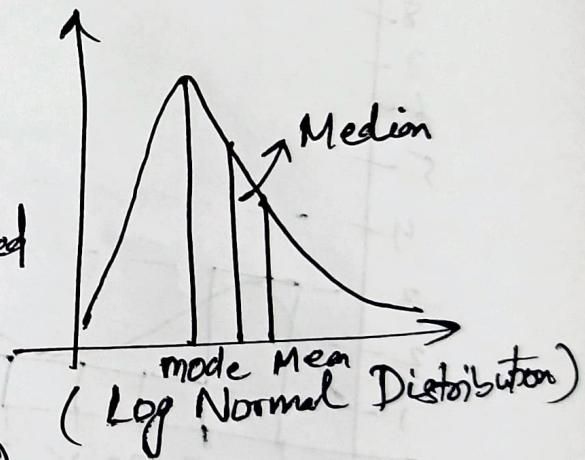
- ① The mean, median, mode all are at perfect centre
 $\rightarrow (\text{Mean} = \text{Median} = \text{Mode})$



Right Skewed data

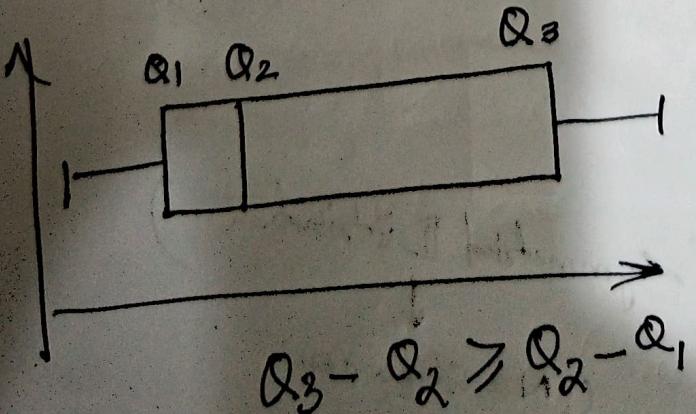


→ Positive Skewed

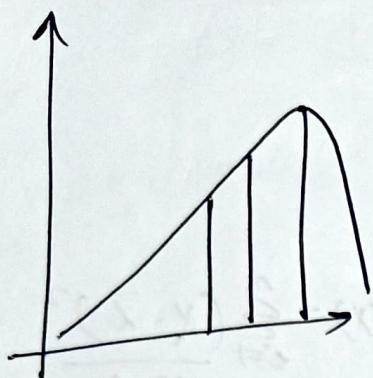


(Relationship b/w mean, median, mode)

$\boxed{\text{Mean} > \text{Median} > \text{Mode}}$

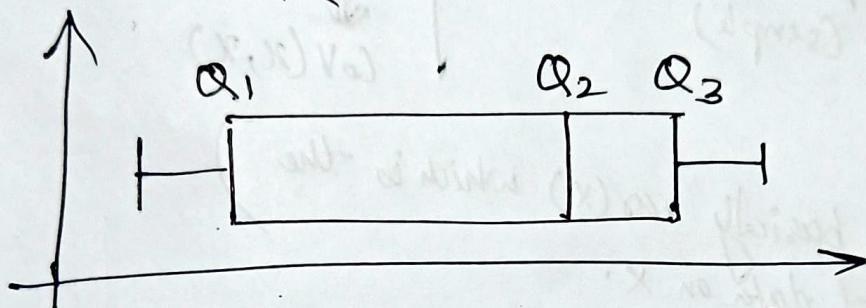


Left-skewed distribution



\Rightarrow Negative skewed

Mean $<$ Median $<$ Mode



$$Q_2 - Q_1 \geq Q_3 - Q_2$$

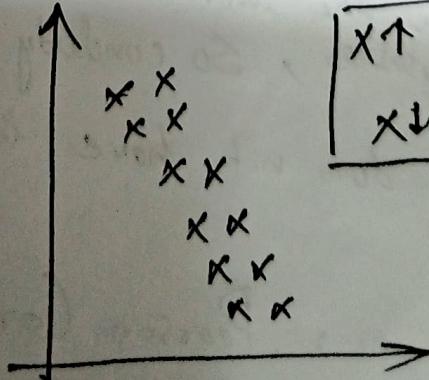
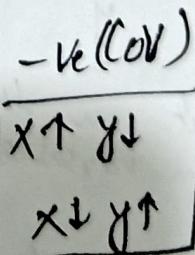
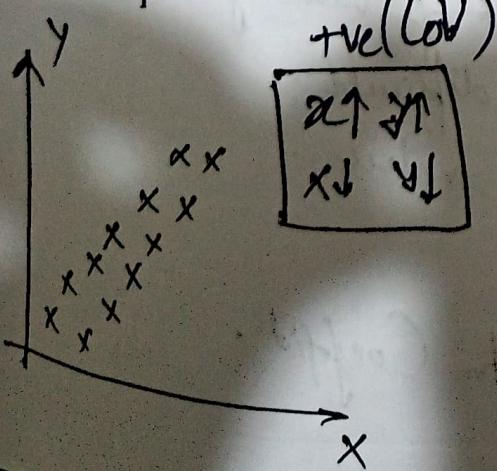
Covariance & Correlation

X	Y
2	3
4	5
6	7
8	9

[Relationship b/w X and Y]

can be

- $X \uparrow Y \uparrow$
- $X \downarrow Y \uparrow$
- $X \downarrow Y \downarrow$
- $X \uparrow Y \downarrow$



Covariance

$$\text{Cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)}$$

x_i = data points

\bar{x} = mean (sample)

y_i = data points

\bar{y} = mean (sample)

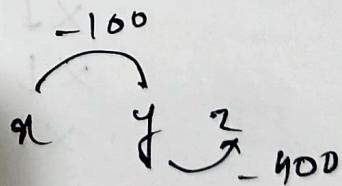
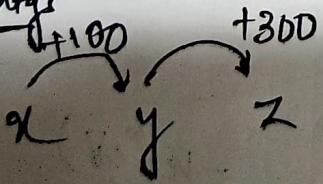
$$\left| \begin{array}{l} \text{Var}(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \\ \Downarrow \\ \text{Cov}(x, x) \end{array} \right.$$

* $(\text{Cov}(x, x))$ is basically $\text{Var}(x)$ which is the spread of data on x .

Advantages of Covariance

- ① find out the relationship b/w x and y .
- ②

Disadvantages



- * No limit on the increment or decrement of value of random variables, so concluding is difficult.
- * Covariance does not have limit values.

To fix this we use Pearson Correlation Coeff.

Pearson Correlation Coefficient [-1 to 1]

$$\rho_{x,y} = \frac{\text{Cov}(x,y)}{\sigma_x \sigma_y}$$

σ_x : Std of x
 σ_y : Std of y

(for linear relationships) or approx (Normal distribution) because of this Cov is limited

- ① The more the value towards +1, the more +vely correlated it is.
- ② The more the value towards -1, the more -vely correlated it is.

Spearman Rank Correlation [-1 to 1]

$$\rho_s = \frac{\text{Cov}(R(x), R(y))}{\sigma_{R(x)} \times \sigma_{R(y)}}$$

$R(x)$ = Rank of x
 $R(y)$ = Rank of y .

eg

x	y	$R(x)$	$R(y)$
1	2	5	5
3	4	4	4
5	6	3	3
7	8	2	1
0	7	6	2
8	1	1	6

(preferrd)

for data Not
 ① Normally distibl
 ② has Outliers
 ③ Not strictly linear

Implementation

① Feature Selection

Size of flat ↑
 (+ve)

No of rooms ↑
 (+ve)

Location ↑
 (+ve)

No of people stay
 ≈ 0 (No relation)

Leisure C-ve

Price ↑

(choose imp features)

(we can delete feature which do not affect the output)

* Correlation and Covariance Relationship

Covariance : tells you how two random variables are connected to each other.
but does not give you a measure of strength.

Correlation : makes the relationship as standardized relationship which is a simple scale from -1 to 1.
It tells you the strength of a relationship.

Kendall Correlation Coeffic

It also tells the strength and direction, but it maps the value on basis of its agreement or disagreement.

formula

$$\tau = \frac{C - D}{\frac{1}{2} n(n-1)}$$

C-Concordant : if ranks of all pairs agree with each other.

D-Discordant, if ranks of all the pairs with each other.

If most pairs are in same order = 1

If most pairs are in opposite order = -1

If most pairs are in random = 0

- * Use "Pearson" if you care about the actual values and assume a straight line relationship.
- * Use "Spearman" if you care about the ranks of the values
- * Use "Kendall" if you want to measure the agreement or disagreement of the order of data points

Syntax Covariance
`df.cov()`

Corr

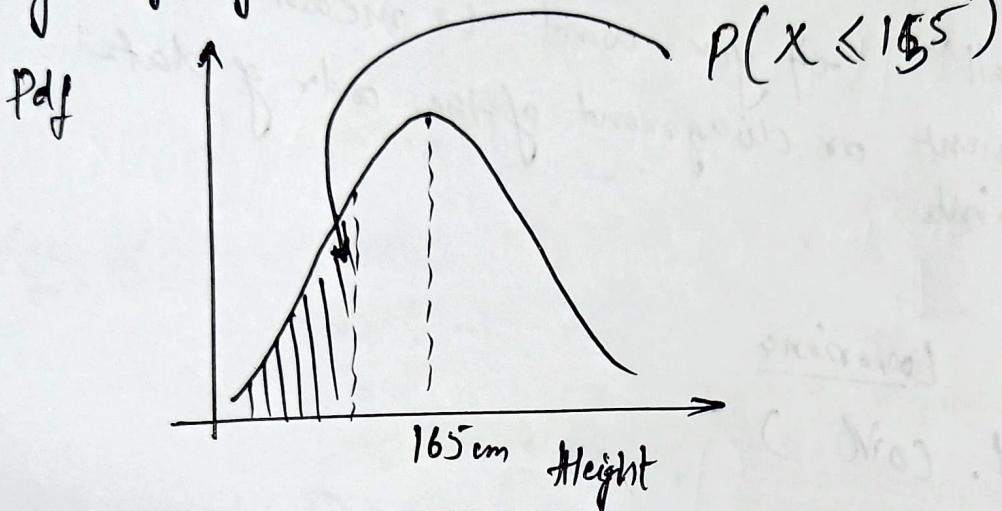
- ① `df.corr(method='pearson', numeric_only=True)`
- ② `df.corr(method='spearman', numeric_only=True)`
- ③ `df.corr(method='kendall', numeric_only=True)`

Probability Distribution function

(1) Continuous Random Variable

Probability Density function

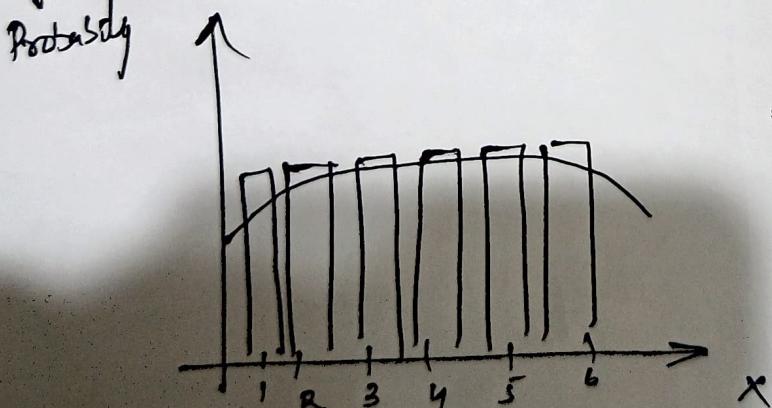
e.g.: Height of student in Classroom.



(2) Probability Mass function

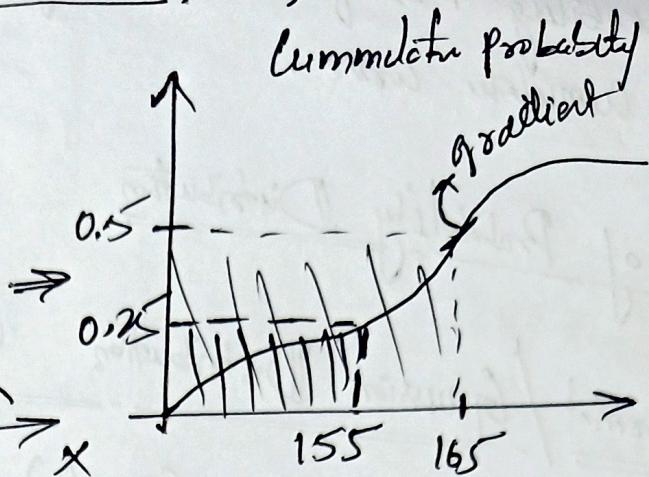
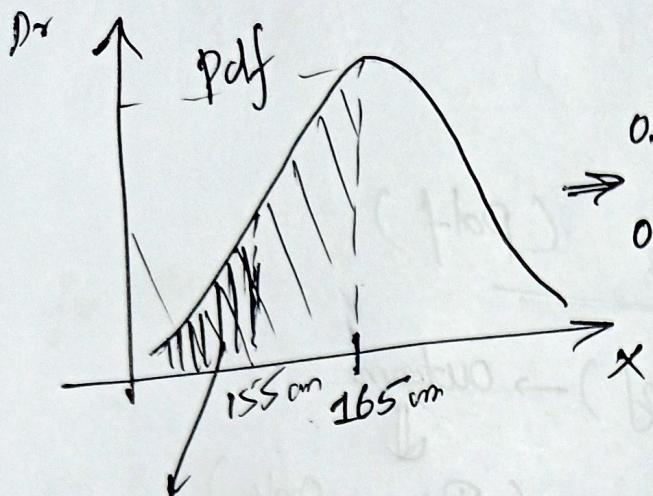
for discrete Random Variable

e.g.: Rolling a dice $\{1, 2, 3, 4, 5, 6\}$

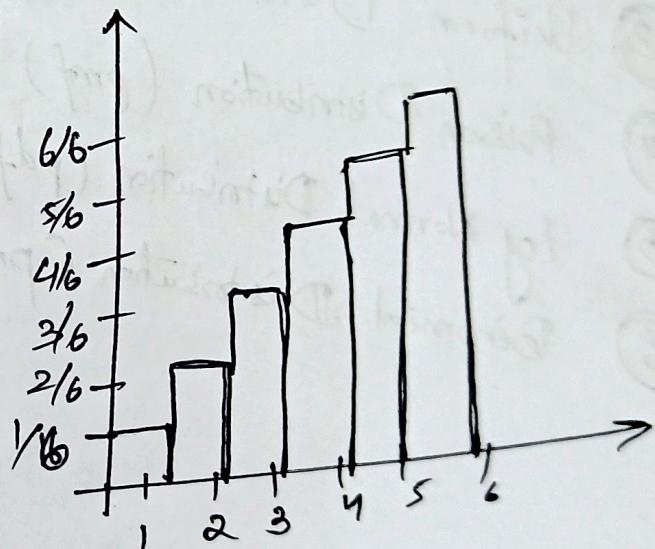
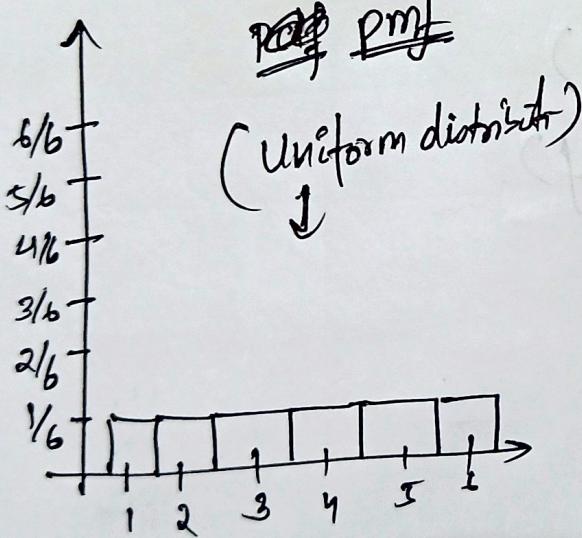


$$\begin{aligned}
 P(X \leq 4) &= P(X=1) + P(X=2) + P(X=3) \\
 &\quad + P(X=4) \\
 &= \frac{4}{6} = \frac{2}{3}
 \end{aligned}$$

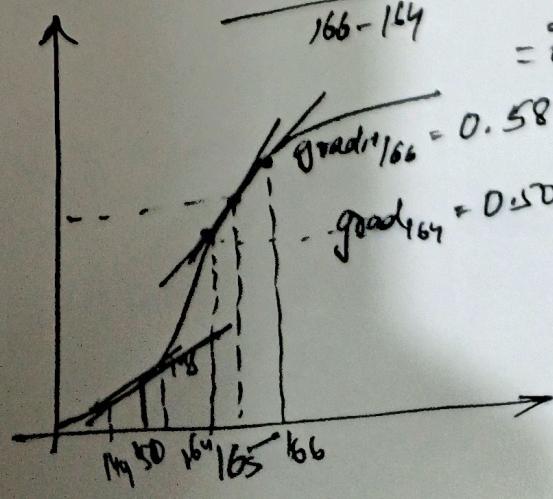
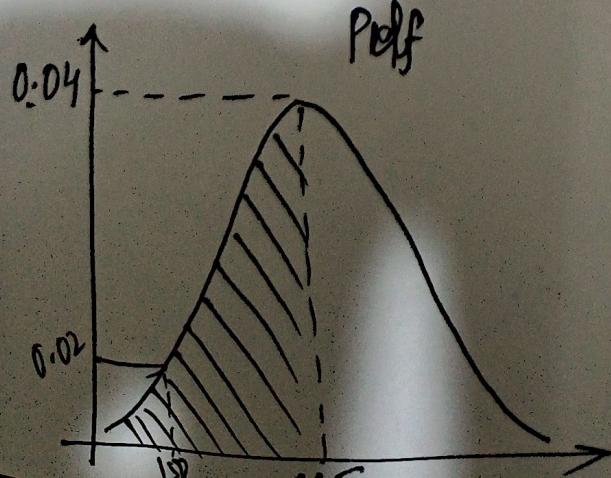
(ii) Commutative Distribution function (cdf)



for dice (Discrete Random Variable) cdf



Distribution of Continuous Random Variable



$$\frac{\text{grad}_{165} - \text{grad}_{164}}{166 - 164} = \frac{0.08}{2} = 0.04$$

Probability Mass function is the graph value of
Cumulative curve.

Types of Probability Distribution

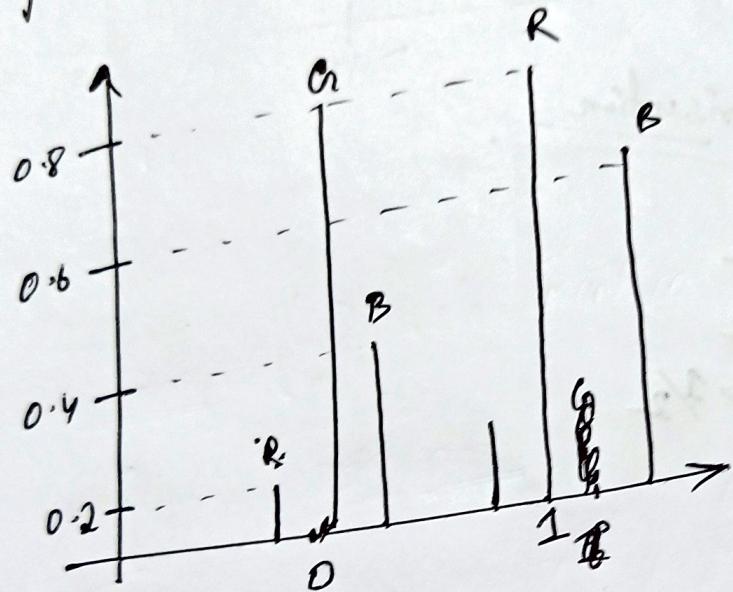
- ① Normal / Gaussian Distribution (pdf)
- ② Binomial Distribution (pmf) → outcome
 ↓
 (Binary Only)
- ③ Uniform Distribution (pmf)
- ④ Poisson Distribution (pmf)
- ⑤ Log Normal Distribution (pdf)
- ⑥ Binomial Distribution (pmf)

Bernoulli Distribution

two outcomes; i.e. outcomes are binary

success (P) failure $(q = 1 - P)$

for discrete random variable, it follows pmf (Probability mass function)



Red $P(x=0) = 0.2$

$$P(x=1) = 0.8$$

Blue $P(x=0) = 0.4$

$$P(x=1) = 0.6$$

Green $P(x=0) = 0.8$

$$P(x=1) = 0.2$$

Parameters:

$$0 \leq p \leq 1$$

$$q = 1 - p$$

$$K = \{0, 1\}$$

pmf

$$\boxed{P^k q^{1-k}}, \quad k \in \{0, 1\}$$

Mean of Bernoulli Distribution

$$E(K) = \sum_{k=0}^K k \cdot p(k)$$

assume
 $P(K=1) = 0.6 \rightarrow P$
 $P(K=0) = 0.4 \rightarrow 1-P = 0.4$

if

$$\sum_{k=0}^K k \cdot p(k)$$

$$\Rightarrow [0 \cdot 0.4 + 1 \cdot 0.6] \\ = 0.6$$

Median of Bernoulli Distribution

$$\text{Median} \begin{cases} 0 & \text{if } P < \frac{1}{2} \\ [0,1] & \text{if } P = \frac{1}{2} \\ 1 & \text{if } P > \frac{1}{2} \end{cases}$$

Variance

$$\text{Var} = P \cdot (1-P)$$

Standard Deviation

$$\text{Std} = \sqrt{P}$$

Binomial Distribution

for discrete random variables.

* Binomial Distribution (n, p)

* Every exp outcome is binary; this experiment is performed for n trials.

* Bernoulli is special case of Binomial where $n=1$.

Binomial ($B(n, p)$)

Parameters: $n \in \{0, 1, 2, \dots\}$
 $p \in \{0, 1\}$ \rightarrow Success probability for each trial.

e.g. Tossing a coin Ten times.

$K \rightarrow$ no of success.

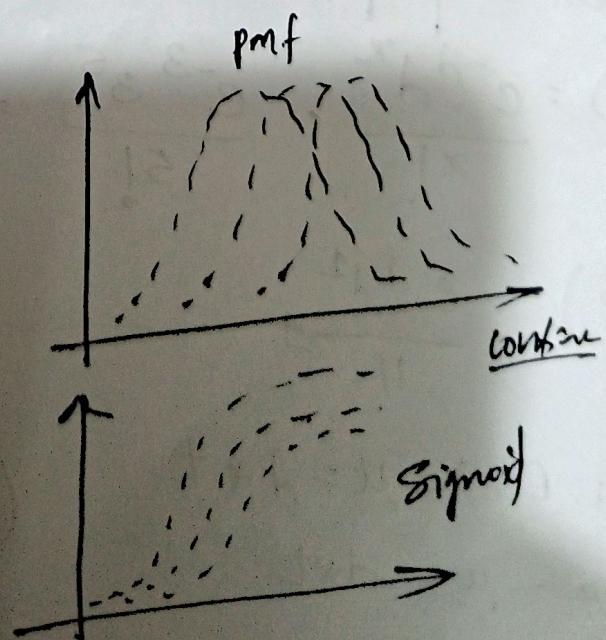
$$\text{pmf} = P(K, n, p) = {}^n C_K p^k (1-p)^{n-k}$$

for $K=0, 1, 2, \dots, n$, where ${}^n C_K = \frac{n!}{k!(n-k)!}$

$$\text{Mean} = np$$

$$\text{Variance} = npq$$

$$\text{std deviat} = \sqrt{npq}$$

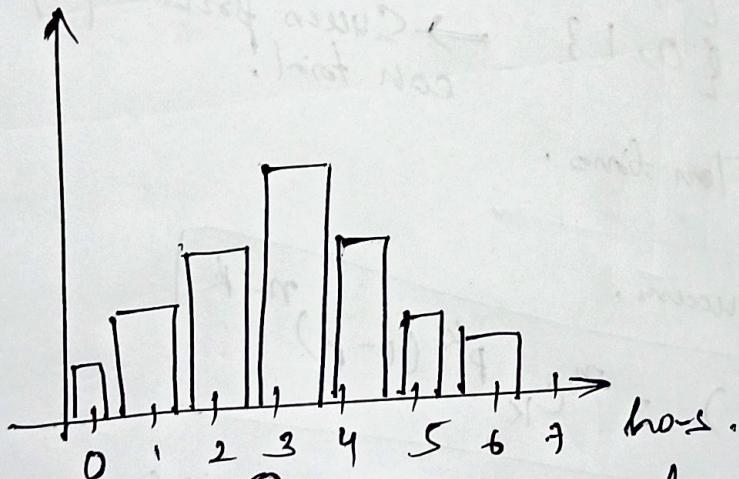


Poisson Distribution

- ① for Discrete Random Variable
- ② Describe the no of events occur in a fix el-time interval.

e.g.: ① No. of people visiting hospital erg hour
 ② No. of people visiting banks erg hour.

pmf



$\lambda = \text{expected no of event to occur at every time interval}$

In context to eq(1).
 When we say $\lambda = 3$, we mean that we expect at least 3 people to visit erg hour.

$$P(X=5) = \frac{e^{-\lambda} \lambda^5}{5!} = \frac{e^{-3} 3^5}{5!} = 0.101 = 10.1\%$$

$$P(X \leq 5) = \frac{e^{-\lambda} \lambda^0}{0!} + \dots - - - + \frac{e^{-\lambda} \lambda^5}{5!}$$

Mean: $E(X) = \mu = \lambda * t$ ($t = \text{time interval}$)

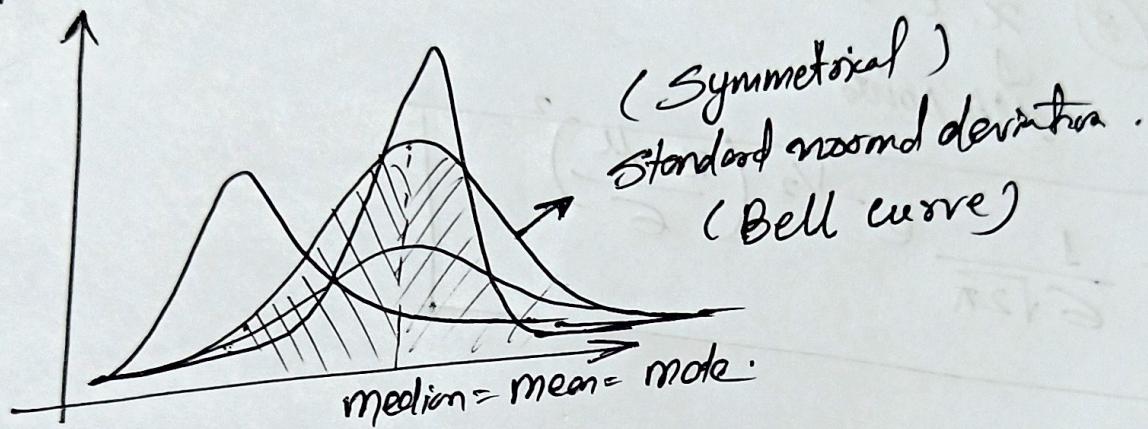
Variance = $\mu = \lambda * t$

Std =

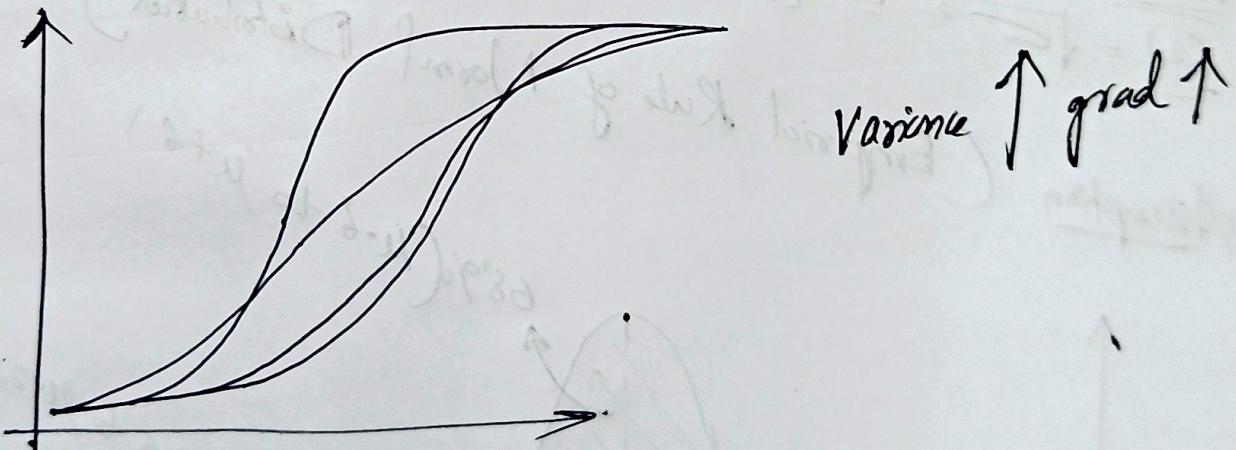
Normal or Gaussian 1D distribution

- * It is for continuous random values.
- * It is for real valued random variable.

Pdf (Probability Density function)



Cumulative Density function



Most of the datasets have this kind of curve

e.g.: Ages

TQ
height
weight
length
width

Notation

$$N(\mu, \sigma^2)$$

- Parameters
- ① Mean (μ) $\in \mathbb{R}$
 - ② Variance (σ^2) $\in \mathbb{R}$
 - ③ $x \in \mathbb{R}$
 - ④ data points

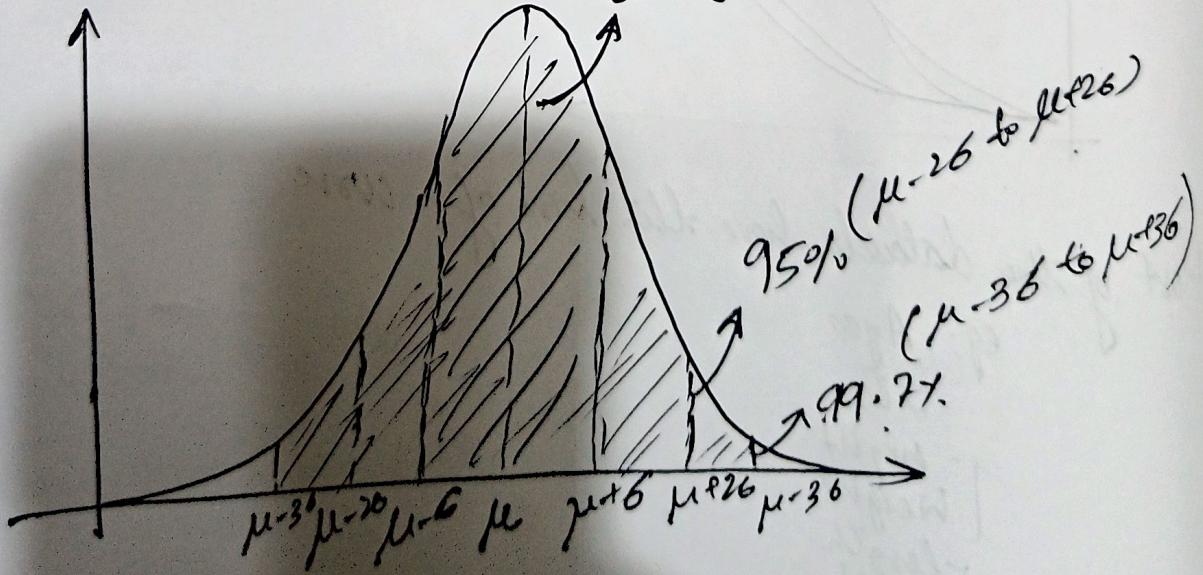
$$\boxed{\text{Pdf} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}$$

$$\text{Mean } (\mu) = \text{Avg}$$

$$\text{Variance} = \sigma^2$$

$$\text{Std} = \sqrt{\sigma^2} = \sigma$$

Assumption (Empirical Rule of Normal Distribution)



$$P(\mu - \sigma \leq x \leq \mu + \sigma) \approx 68\%$$

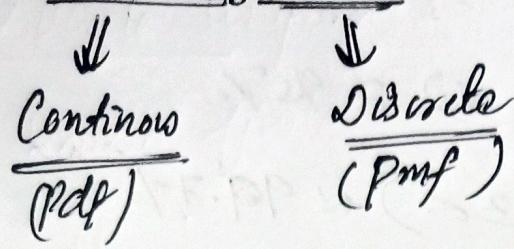
$$P(\mu - 2\sigma \leq x \leq \mu + 2\sigma) \approx 95\%$$

$$P(\mu - 3\sigma \leq x \leq \mu + 3\sigma) \approx 99.7\%$$

Example

- Dataset which contains
- (i) weights of students in the class.
 - (ii) Height of " " " "
 - (iii) PRIS Dataset (sepal width)

Uniform distribution



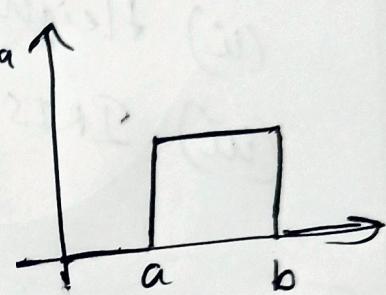
(3) Continuous Uniform (Rectangular distribution)

The bounds are defined by the parameters a and b which are the minimum and maximum values.

Notat: $U(a, b)$

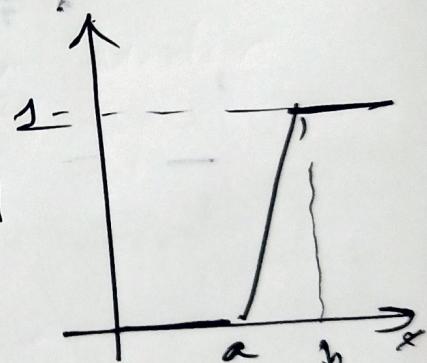
Parameters : ~~$0 < a < b$~~

$$-\infty < a < b \text{ so } a, b$$



$$\text{P.d.f} = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Cdf} = \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } x \in [a, b] \\ 1 & \text{for } x > b \end{cases}$$



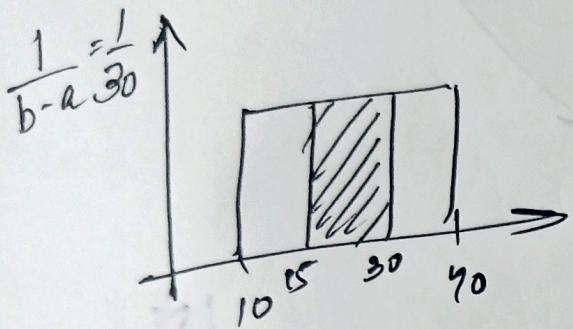
$$\text{Mean} = \frac{1}{2}(a+b)$$

$$\text{Median} = \frac{1}{2}(a+b)$$

$$\text{Variance} = \frac{1}{12}(a+b)^2$$

eg: The number of candies sold daily at a shop is uniformly distributed with a maxm of 40 and a minm of 10.

(i) Prob of daily sales to fall b/w 15 and 30?



$$\begin{aligned}
 P(15 \leq x \leq 30) &= (x_2 - x_1) \times \frac{1}{b-a} \\
 &= (30 - 15) \times \frac{1}{30} \\
 &= \frac{15}{30} = 0.5 \text{ or } 50\%
 \end{aligned}$$

(ii) Prob of sales > 20

$$\begin{aligned}
 P(x \geq 20) &= (40 - 20) \times \frac{1}{30} \\
 &= \frac{20}{30} = 0.66 \text{ or } 66.6\%
 \end{aligned}$$

② Discrete Uniform Distribution
 a known finite no of outcomes equally likely to happen.

e.g. Rolling a die

1, 2, 3, 4, 5, 6

$$\begin{array}{l} a=1 \\ b=6 \end{array} \quad \begin{array}{l} P(x=1) = \frac{1}{6} \\ P(x=2) = \frac{1}{6} \\ \vdots \\ P(x=6) = \frac{1}{6} \end{array}$$

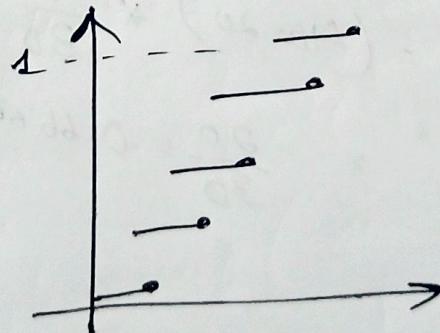
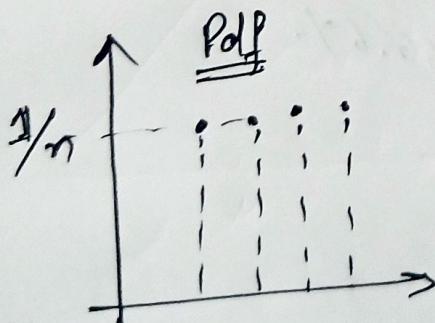
Notation : $U(a, b)$

Param : a, b with $b \geq a$

$$\text{mean} = \frac{a+b}{2} = \text{Median}$$

$$\underline{\text{Pmf}}: \frac{1}{n}$$

$$\underline{\text{CDF}}: \sum_{i=1}^n \frac{1}{n} = 1$$



Standard Normal Distribution and Z-score

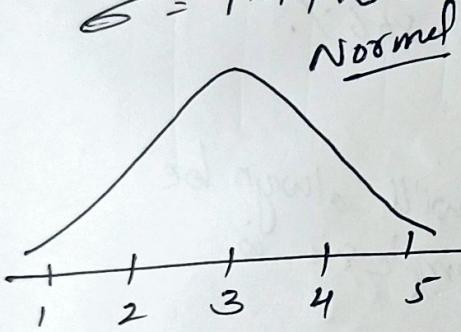
Z-score

Assumij Normal Distro with

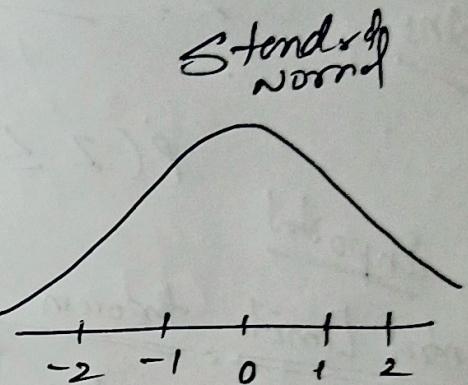
$$x = \{1, 2, 3, 4, 5\}$$

$$\mu = 3$$

$$\sigma = 1.44 \approx 1$$



$$\Rightarrow \begin{matrix} \mu = 0 \\ \sigma = 1 \end{matrix}$$



$$Z\text{-score} = \frac{x_i - \mu}{\sigma}$$

$$= \frac{1-3}{1} = -2$$

$$= \frac{2-3}{1} = -1$$

$$= \frac{3-3}{1} = 0$$

$$= \frac{4-3}{1} = 1$$

$$= \frac{5-3}{1} = 2$$

Use to find how far from the standard deviate does a value fall.

$$\text{const the } P(\alpha < x < \beta)$$

$$P(\quad \leq z \leq \quad)$$

↳ use z-table to calculate
the area under curve.

Q In India the avg IQ is 100, with a $\frac{Std(6^2)}{15} = 15$. What is the % of the population would you expect to have an IQ lower than 85?

Ans $Z = \frac{85 - 100}{15} = -\frac{15}{15} = -1$

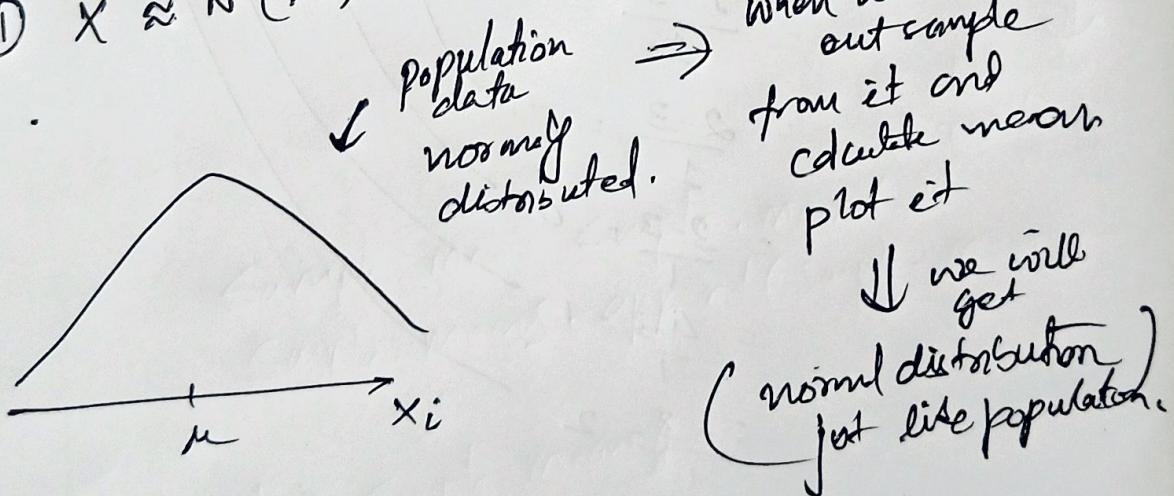
$$P(Z \leq -1) = \frac{0.15866}{15.866\%}$$

** Important

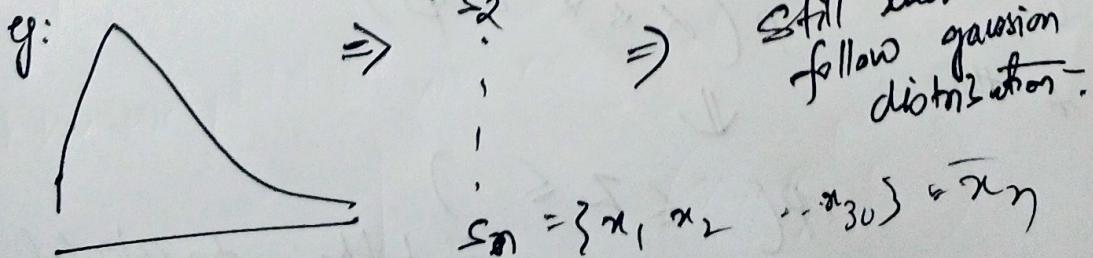
Central Limit theorem

The sampling distribution of the mean will always be normally distributed as long as the sample size is large enough.

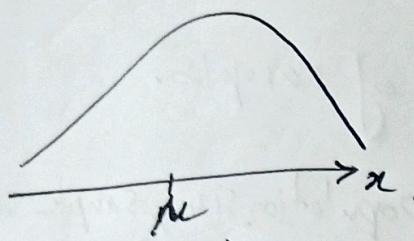
① $X \approx N(\mu, \sigma)$



② $X \notin N(\mu, \sigma)$ $S_1 = \{x_1, x_2, \dots, x_{30}\} \Rightarrow \bar{x}_1$



Interviews



$$X \sim N(\mu, \sigma)$$

σ = pop. std

μ = pop. mean

n = sample size

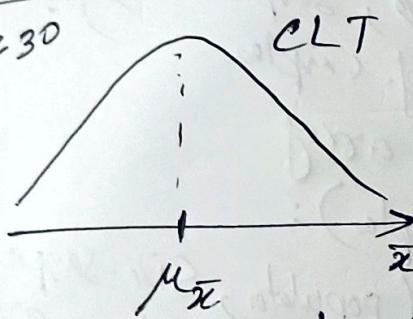
approximately equal to

assume (becan $n > 30$)

↓
Sample distns of mean

assumption

$$\sqrt{n} \geq 30$$



$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

If n can be
any value for gaussian
distribution

Key points

(1) Population distribution

The original population from which samples are drawn
Can have any distribution.

(2) But as the sample size increases the sample mean tends to approach a normal distribution.

(3) Samples must be drawn independently for the same distribution for the CLT to hold.

Code Breakdown

Population-size = 10 000 → a big population

Sample-size = 100

No-of-samples = 1000 → Number of samples.

population = np.random.exponential(population_size, sample_size, No-of-samples)

* generation exponential distribution
to hold CLT.

def calculate_sample_mean(population_size, sample_size, No-of-samples):

 Sample-means = [] # creating array

 for _ in range(No-of-samples):

 sample = np.random.choice(population_size, sample_size, replace=False)

 ↓
(this to have
randomness and
no bias)

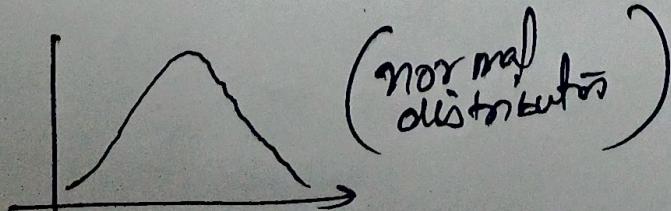
 Sample-mean = sample.mean()

 Sample-means.append(sample.mean())

return Sample-means

Sample-means = calculate_sample_mean(population_size, sample_size, No-of-samples)

When we plot this we nearly get



"Inferential Statistics" Estimate.

It is an observed numerical value used to estimate an unknown population parameter.

Two types

① Point Estimator: Single numerical value used to estimate the unknown population parameter.

② Interval Estimator: Range of values used to estimate the unknown population parameter.

* Interval Estimate of population parameters are called "Confidence Intervals".

