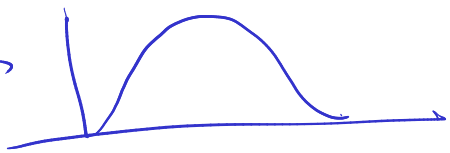
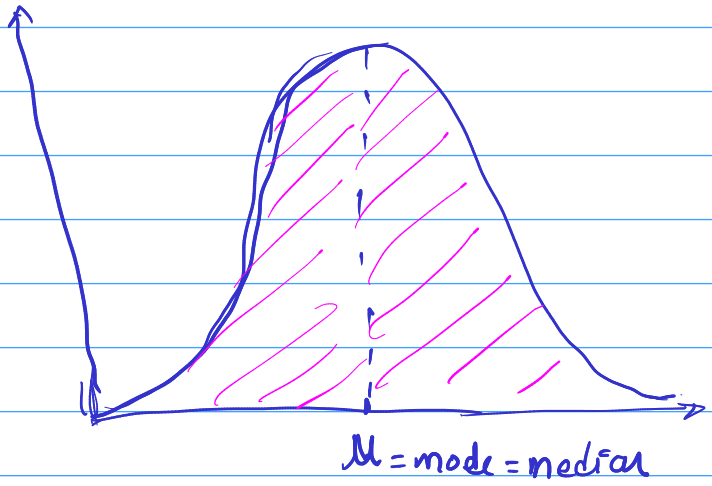
PDF →

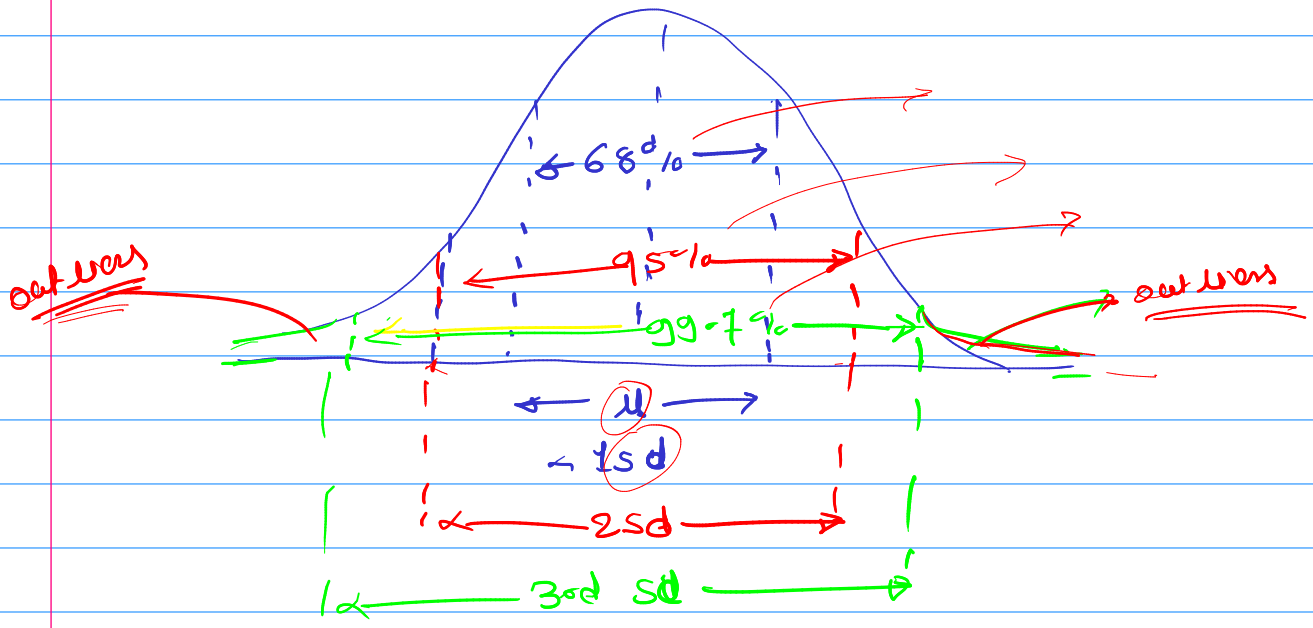


① Normal Distribution (Gaussian Distribution)

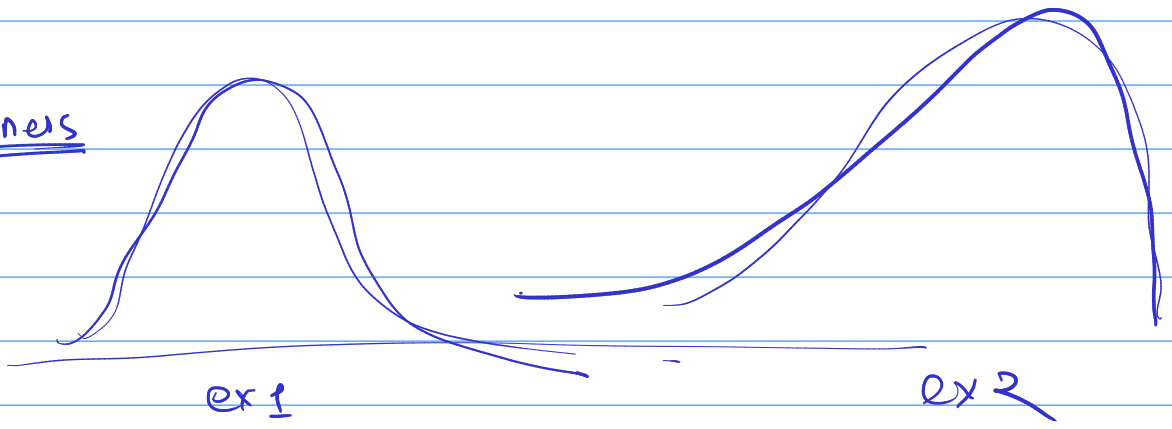


1) Bell curve
2) left side & right is symmetrical
3) mean = mode = median
4) for continuous variable (

5) It follow the empirical Rule (68% 95, 99.7)

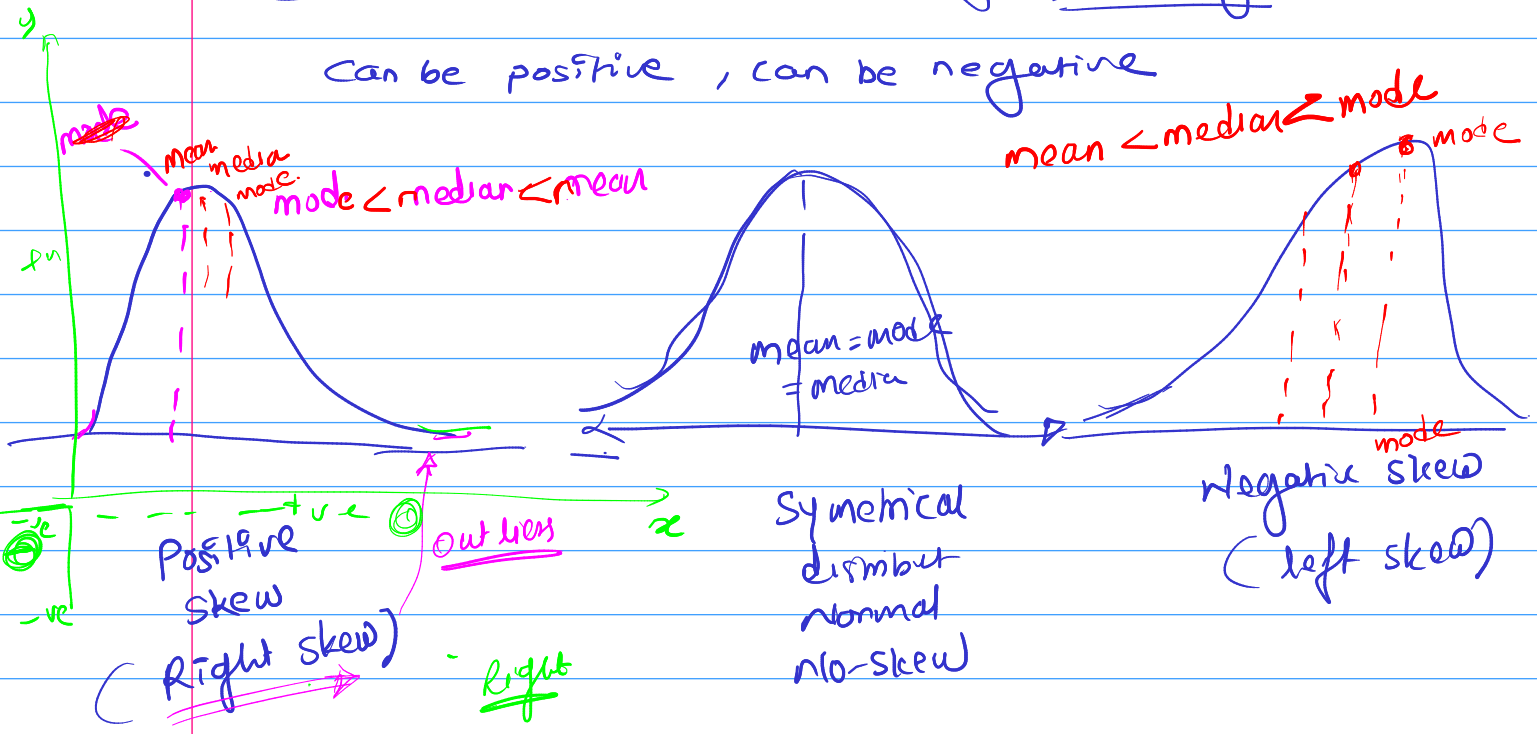


Skewness

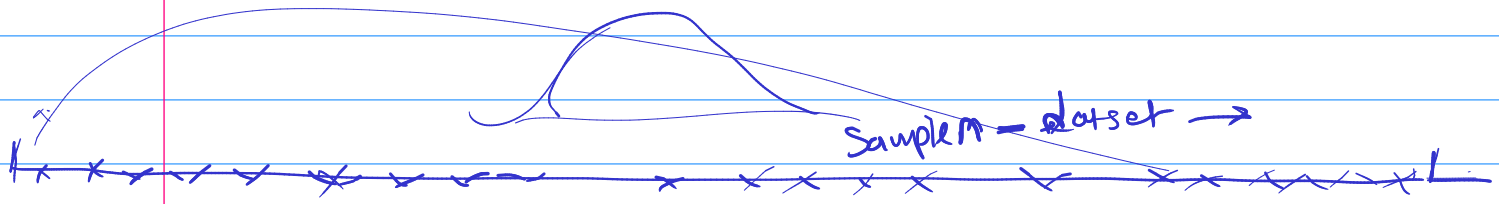


Skewness → is a measurement of asymmetry

Can be positive, can be negative

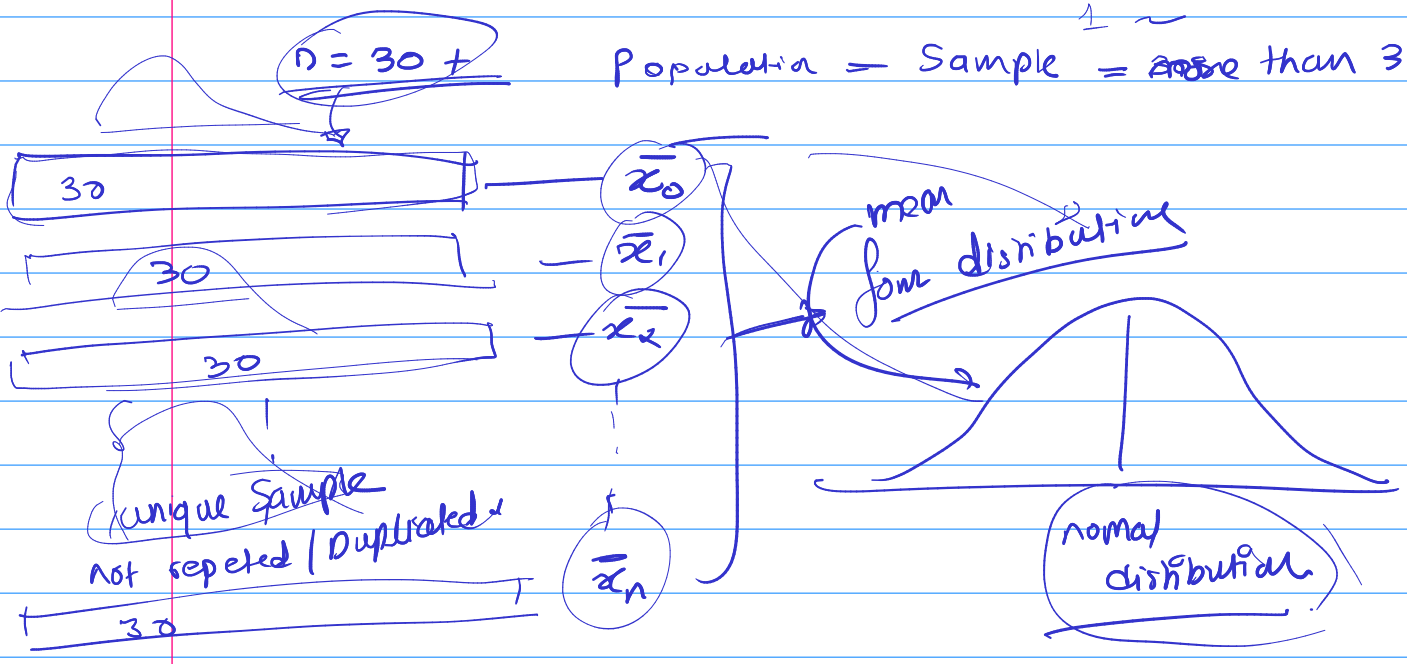


- use for outlier detection also
- Hypothesis tesing
- Central limit theorem



n = 30 +

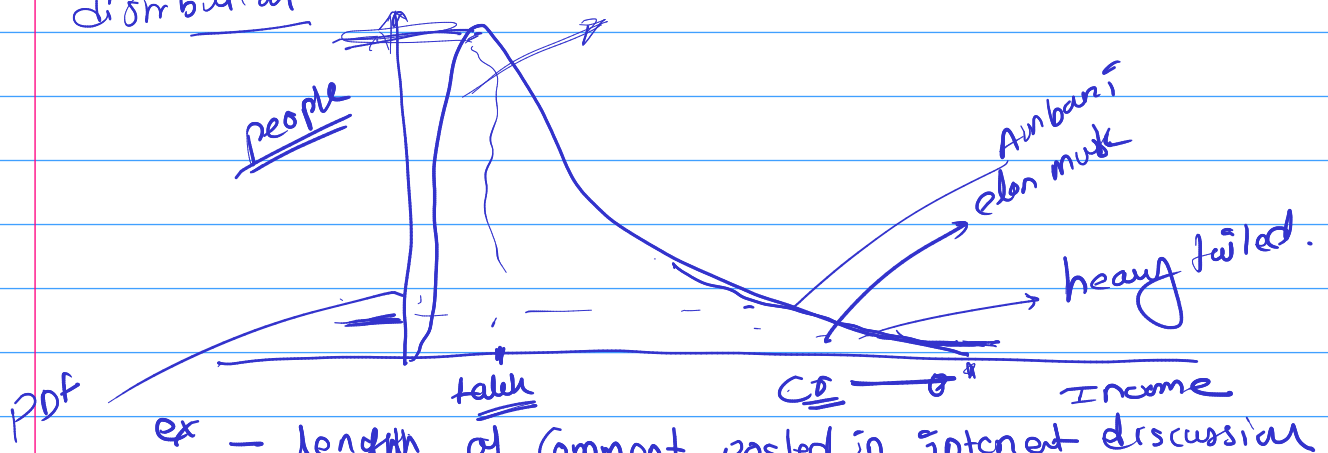
Population = Sample = more than 30 Sample



Central limit theorem in prob.

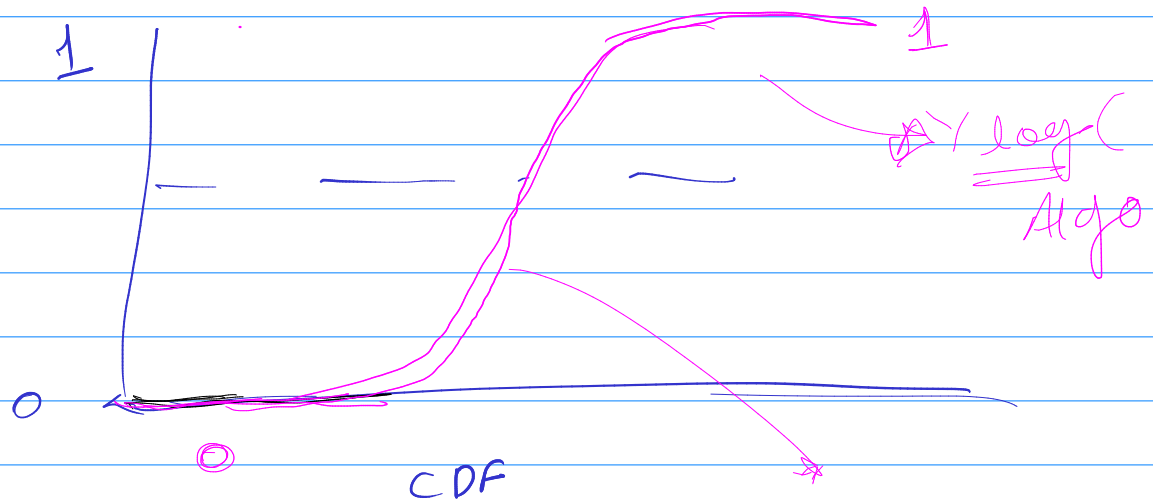
2) Log-Normal Distribution

— is a heavy tailed Continious probabability distribution



ex — length of Comment posted in internet discussion

— Income pareto — 97-99% log normal

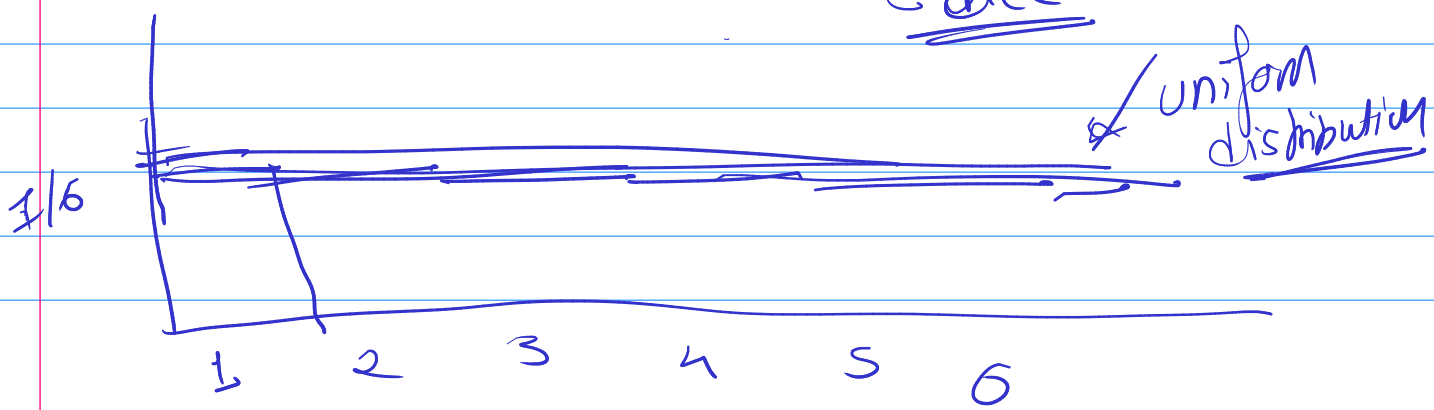


x = log dist

y = ln(x)

(3) Uniform distribution (Descrete values)

↳ dice



(4) Bernoulli Distribution (only 2 outcome)

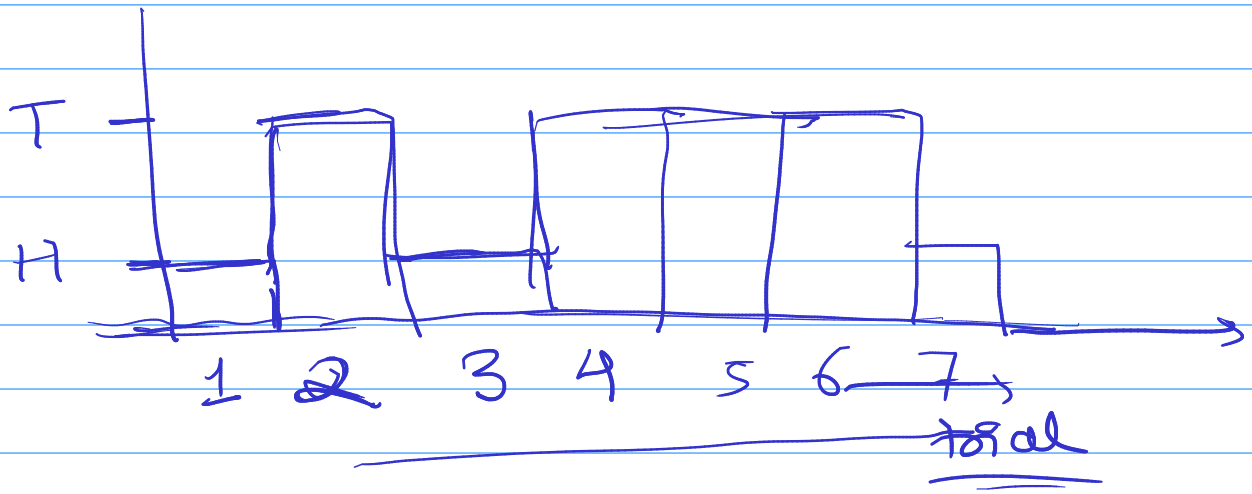
(True / false) (Yes / No) (0, 1)

(discrete)

Head / Tail

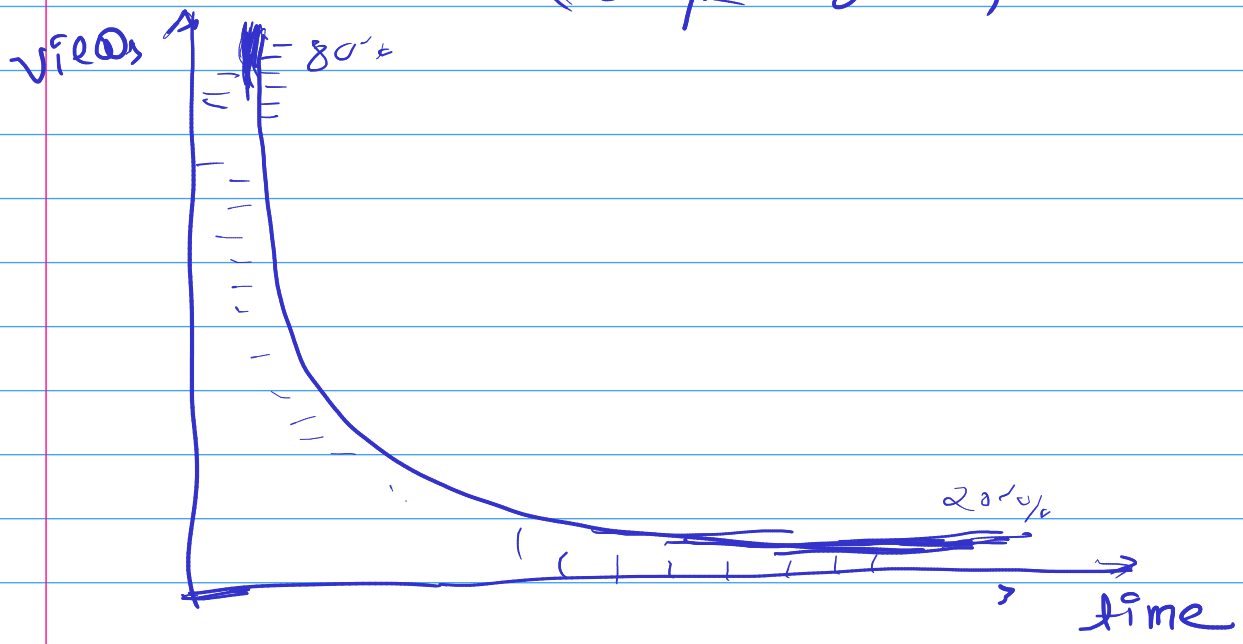
(3) Binomial Distribution → are the sequence of bernoulies.

→ no. of Bernoulies



⑥ Pareto Distribution (Power law of distribution)

(80/20 rule)

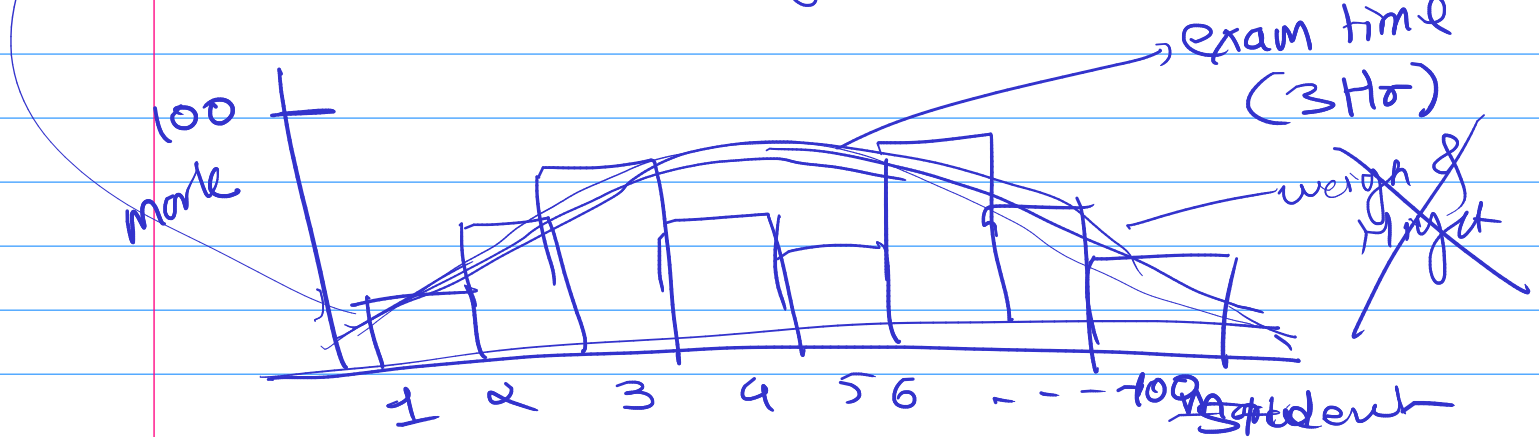


youtube channel — view

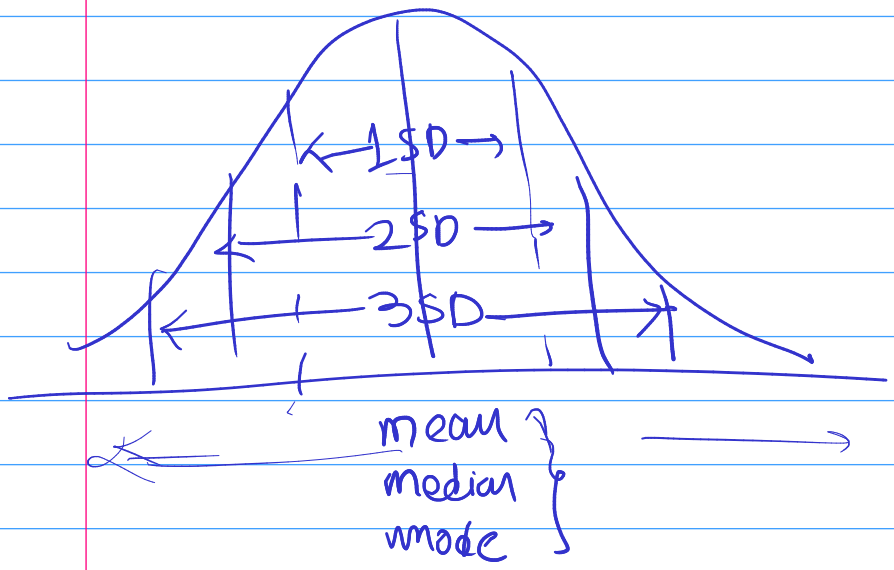
⑦ Poisson Distribution

→ It occurs within a specific interval / time

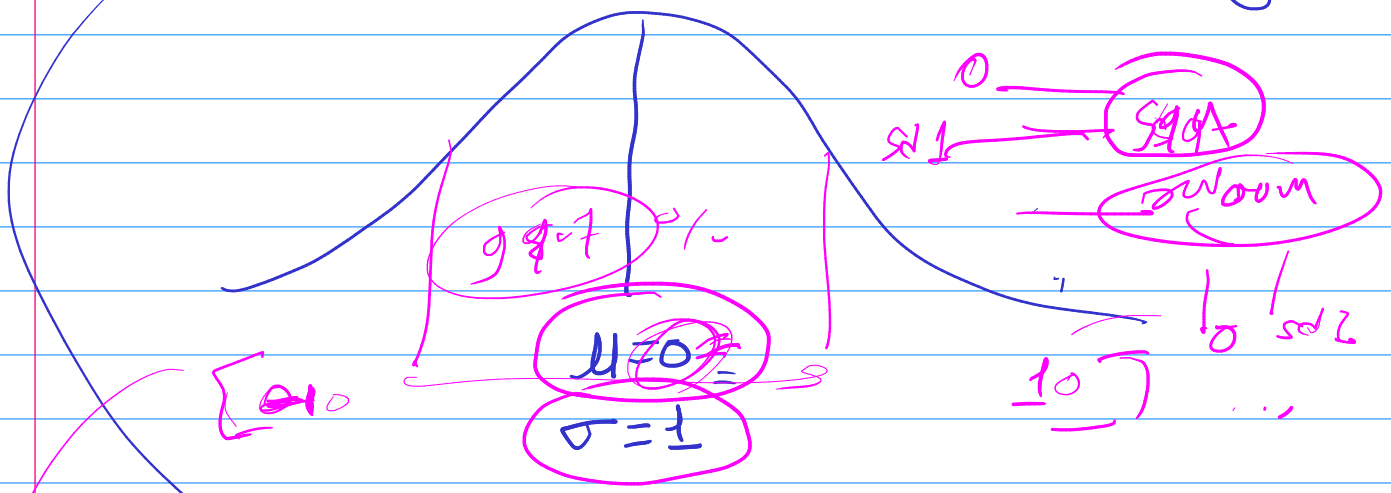
ex - exam mark of the student



⑥ Normal distribution



Standard Normal Distribution (scaling)



Z-test (after applying Z test on normal distribution we get standard normal distribution)

Standerdization → Normalization (min-max Scaler)

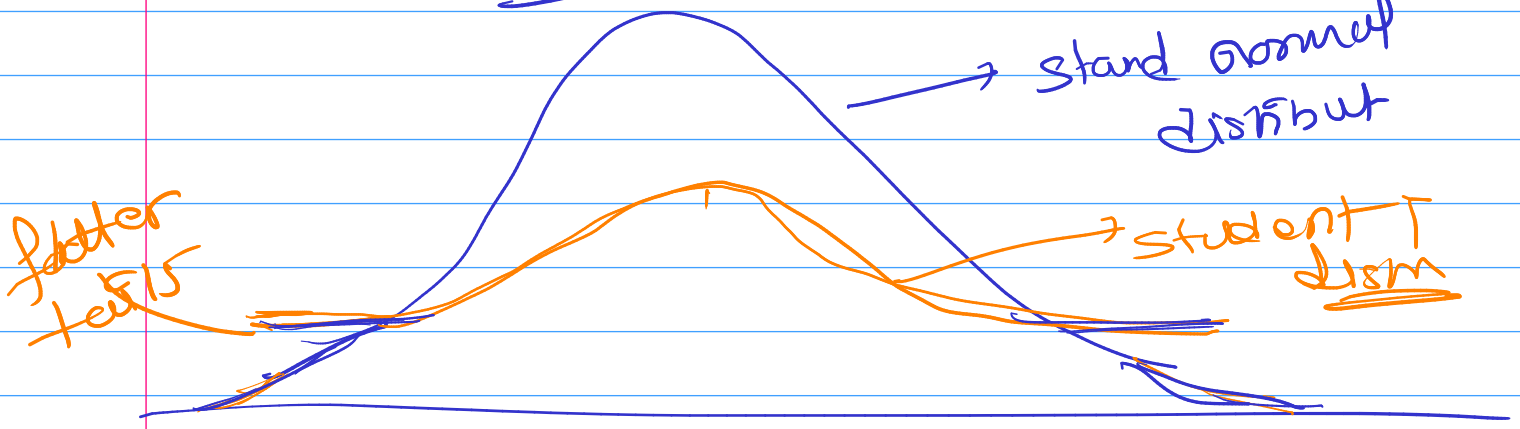
Scaling

start - end

1/255 → 0 - 255 (min, max)

(8) Student T- distribution

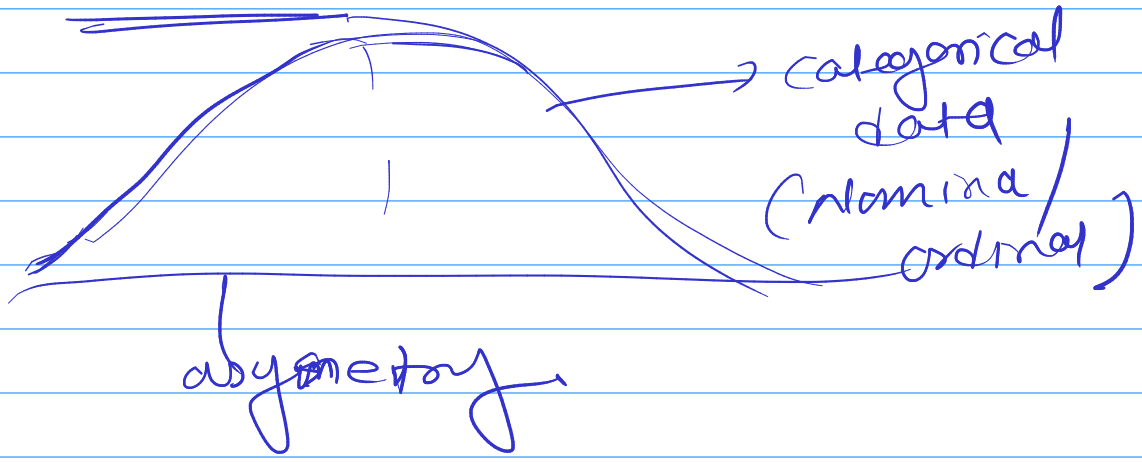
sample size less than (30)



Normal T- test = greater than 30

(9) Chi-square (Categorical Variable)

n — 30 +



Probability = Probability is a measure of the likelihood of the event.

$$\text{Prob} = \frac{\#\text{ no. of way event can occur}}{\#\text{ total no. of possible outcome}}$$

eg - getting 6 after Rolling dice

$$= \frac{1}{6}$$

mutual Exclusive Event / Non-mutual exclusive

Cannot occur at same time

(Rolling dice)

- getting 2 = 1/6
- getting 6 = 1/6

Coin → H / T

Can occur at same time

eg = getting queen Card & Heart Card.

Total = 52 Cards

$P(Q) = 4/52$

$P(♡) = 13/52$

$P(Q \text{ of } ♡) = 1/52$

Addition Rule

$P(A \text{ or } B) = P(A) + P(B)$ — mutual

$P(A \cap B)$ — intercept

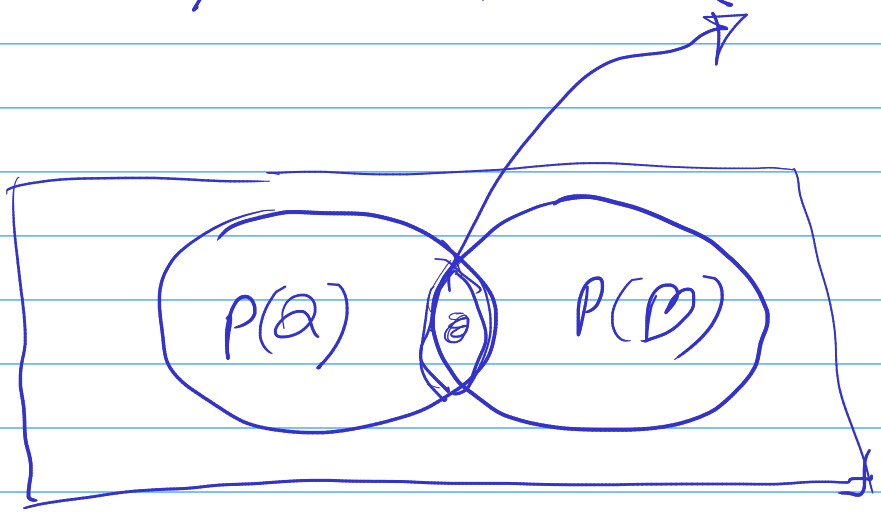
dice =

$P(1) \text{ or } P(3) = 1/6 + 1/6$

$P(A, B) = P(A) + P(B) - P(A \text{ or } B)$

$P(A \cup B) = P(\heartsuit) + P(\diamondsuit) - P(\heartsuit \& \diamondsuit)$

Union



2) multiplication Rule

Independent

all events are independent

Rolling dice

1, 2, 3, 4, ... Independent

Dependent

all events is dependent.

Bag 3 Red ✓
2 greed

total = 5 Ball

P(Red) = 3/5 — 1st event

P(green) = 2/4 → 2nd event

P(Red) = 2/3 → 3rd event

P(green/Red)

P(Red/green)

Naive Bayes = Conditional Probability

Independent

P(A and B) = P(A) * P(B)

Coin = (H, T)

getting

$= 1/2 * 1/2$

→ Dependent event

$$P(Q \& \heartsuit) = P(Q) * P(\heartsuit / Q)$$

Bays Rule

$$P(B) * P(A/B) = P(A) * P(B/A)$$

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

Permutation & Combination

ex ⇒ School trip → chocolate factory visit

6 — dairy milk, 5 Star, eqwiler, munch, mango, Cadbury silk

Permutation → 100 student → factory visit

6 choice — sequence —

— ① ② ③ ④ ⑤ ⑥

= star, cadbury silk, mango, dairy munch

— munch,

100 student = different Combination

500 student = ? 500 Combination

max Combination ? = $^{n}P_{r} = \frac{n!}{(n-r)!}$

total choice → $^{6}C_{3} = \frac{6!}{(6-3)!} = 120$ combination

3 selection

max

Combination ⟹ only 1 combination required

ex = dairy milk, mango, munch

||

same as

3 permutation but 1 combination

ex 2 = mango, munch, dairy milk

||

same

ex = munch, dairy milk mango

- repeated value not count even the order is change

$${}^{n}C_{r} = {}^{6}C_{3} = \frac{n!}{r!(n-r)!}$$

$$= \frac{6!}{3!(6-3)!} = 20 \text{ unique combination}$$

Inferential Stats