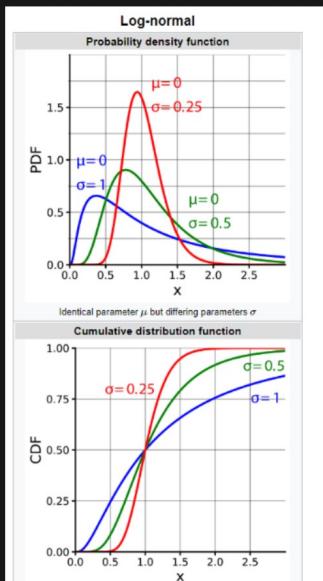


Agenda

Different types of Distribution

- ① Normal / Gaussian Distribution ✓
- ② Standard Normal Distribution ✓
- ③ Log Normal Distribution ✓
- ④ Power Law Distribution ✓
- ⑤ Bernoulli Distribution ✓
- ⑥ Binomial Distribution ✓
- ⑦ Poisson Distribution ✓
- ⑧ Uniform Distribution
 - Discrete ✓
 - Continuous ✓
- ⑨ Exponential Distribution ✓
- ⑩ F distribution ✗
- ⑪ Chi Square distribution. } ✓ } ANNOVA
- ⑫ Hypothesis Testing } ✓ .

① Log Normal Distribution {Continuous Random Variable}.

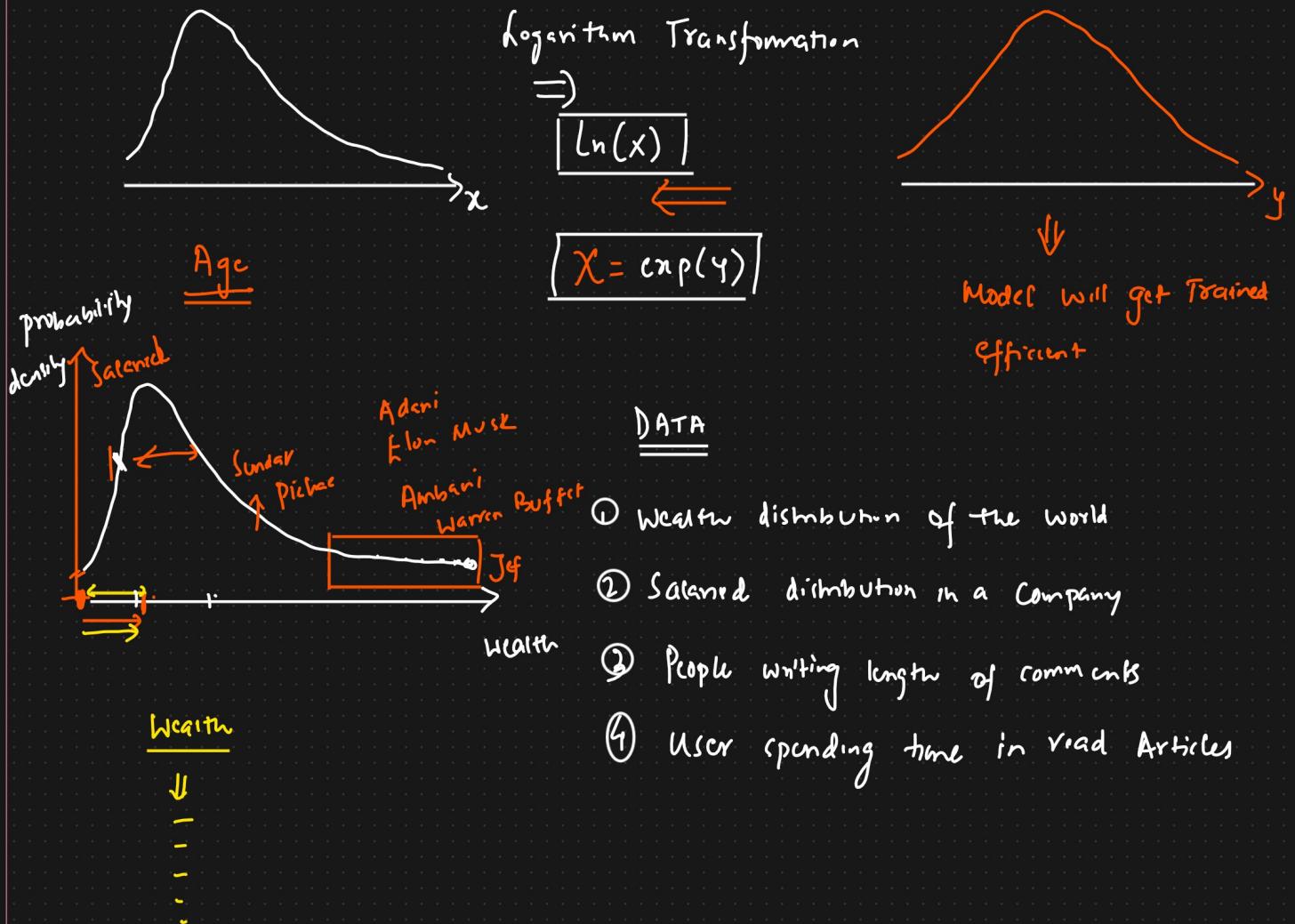


In probability theory, a log-normal (or lognormal) distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then $Y = \ln(X)$ has a normal distribution. Equivalently, if Y has a normal distribution, then the exponential function of Y , $X = \exp(Y)$, has a log-normal distribution.

$$X \sim \text{lognormal}(\mu, \sigma^2)$$

$$Y \sim \ln(x) \Rightarrow \text{Normal Distribution}(\mu, \sigma^2).$$

$\ln = \text{natural log (log e)}$



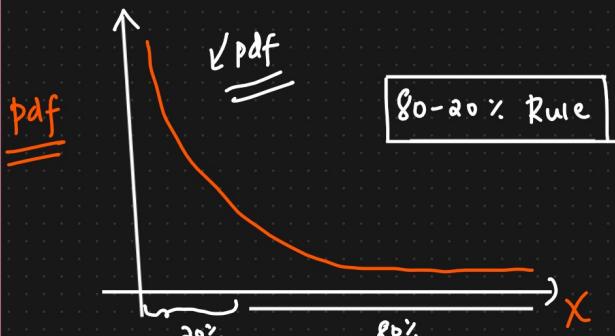
Notation $\text{logNormal}(\mu, \sigma^2)$

Parameters $\mu \in (-\infty, +\infty)$

$$\sigma > 0$$

Pdf $\therefore \frac{1}{x \sigma \sqrt{2\pi}} \exp \left(-\frac{(\ln x - \mu)^2}{2\sigma^2} \right)$

④ Power Law Distribution [Continuous Random Variable]



In statistics, a power law is a functional relationship between two quantities, where a relative change in one quantity results in a proportional relative change in the other quantity, independent of the initial size of those quantities: one quantity varies as a power of another. For instance, considering the area of a square in terms of the length of its side, if the length is doubled, the area is multiplied by a factor of four

Eg: IPL Games

RCB

Project

1A 1PM 1BA

5-6 Sl=

- ① 20% of the team is responsible for winning 80% of the match
- ② 80% of the sales in Amazon is derived from 20% of the products.
- ③ 80% of the wealth is distributed among 20% of the people.
- ④ 80% of the projection complete by 20% of team.

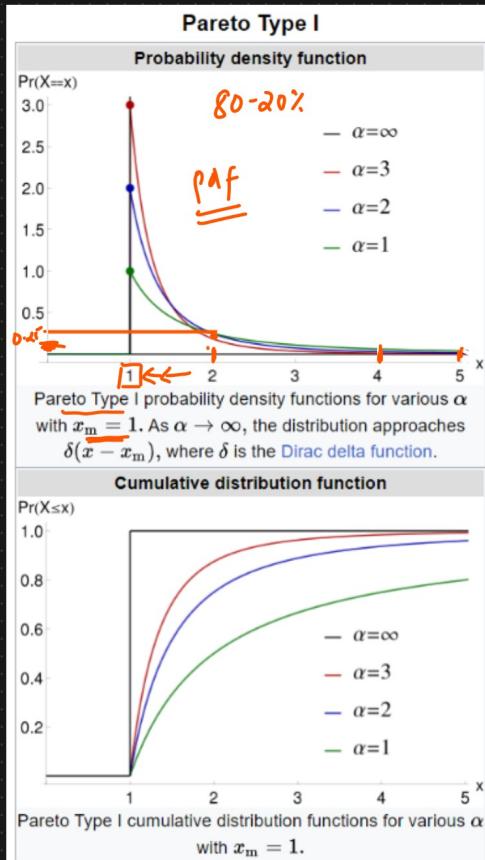
Types of Power Law Distribution

- ① Pareto Distribution
- ② Exponential Distribution

isko net pei search kr liyo or formula yaad rkhnے ki jrurt nii hai

① Pareto Distribution

[Continuous Random Variable]



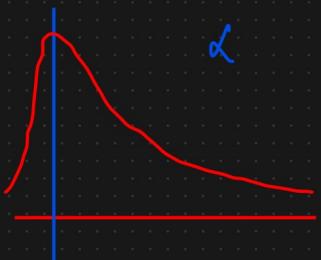
If X is a random variable with a Pareto (Type I) distribution, then the probability that X is greater than some number x , i.e. the survival function (also called tail function), is given by

$$\text{pdf} = \bar{F}(x) = \Pr(X > x) = \begin{cases} \left(\frac{x_m}{x}\right)^\alpha & x \geq x_m, \\ 1 & x < x_m, \end{cases}$$

$$\Pr(X > 4) = \left(\frac{1}{4}\right)^2 = \boxed{\frac{1}{16}}$$

$$\Pr(X = 4) = \left(\frac{1}{2}\right)^2 = \boxed{\frac{1}{4}} = 0.25$$

$$\left(\frac{1}{4}\right)^2 = \boxed{1}$$

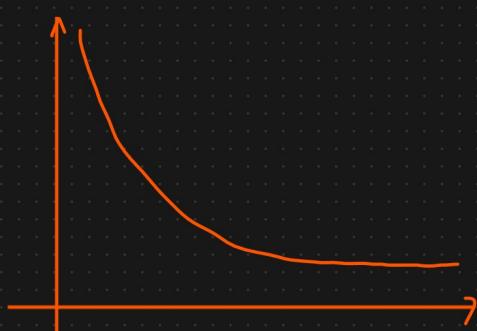
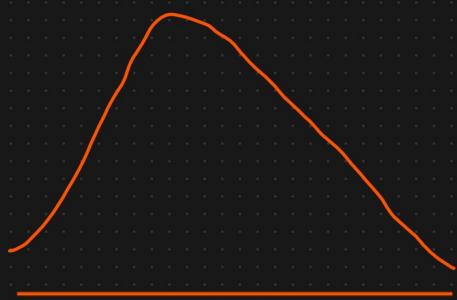


Box Cox

Transformation



f



⑧ Bernoulli Distribution \doteq Outcome of the process is binary {1,0}

Tossing a ^{Fair} Coin $\{ \text{Success, Failure} \}$.

I think it should be PMF bcz it is a discrete random variable or it may be Probability distribution function

$$P_f(H) = 0.5 \Rightarrow P // \Rightarrow P+q = 1$$

$$P_f(T) = 0.5 \Rightarrow 1-P = q //$$

$$\begin{aligned} \text{pdf} &= p^k (1-p)^{1-k} \\ &= p(1-p)^0 \\ &= p // \end{aligned}$$

Graph wagerah net pei dkh liyo

$$K \{ 1, 0 \} \quad \begin{cases} p & \text{if } k=1 \\ 1-p=q & \text{if } k=0 \end{cases}$$

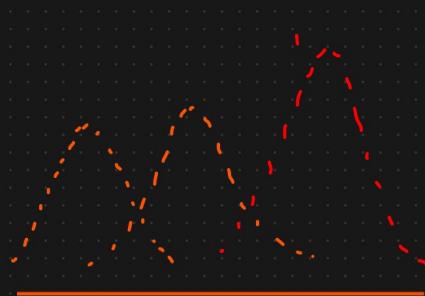
pmf

⑨ Binomial Distribution

\rightarrow Combination of multiple Bernoulli Distribution

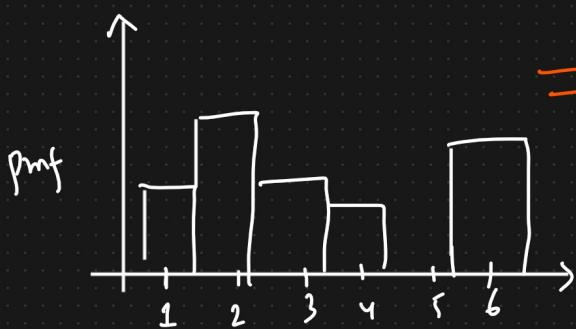
$$n, p = K =$$

$$\text{pmf} = n C_k p^k (1-p)^{n-k}$$



any definition or graph can be seen in wikipedia. So must refer there too

⑩ Poisson Distribution



No. of people visiting bank every hour

$\Rightarrow \lambda = 3$ \Rightarrow Expected no. of people to come at that specific time interval

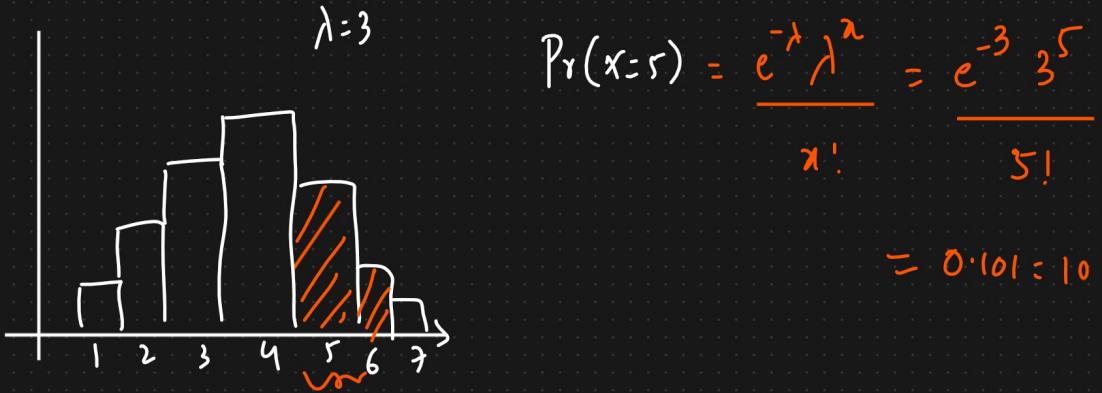
$$\text{pmf} \quad \Pr(X=5) = \frac{e^{-\lambda} \lambda^x}{x!}$$

pdf

Probability density
 $f(x)$

pmf

Probability



$$\begin{aligned}\Pr(X=5 \text{ or } 6) &= \Pr(X=5) + \Pr(X=6) \\ &= \frac{e^{-\lambda} \lambda^5}{5!} + \frac{e^{-\lambda} \lambda^6}{6!}\end{aligned}$$

④ Uniform Distribution

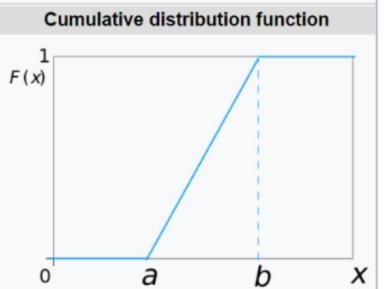
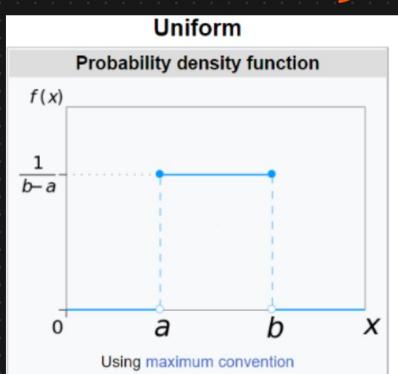
① Continuous Uniform Distribution (pdf)

② Discrete Uniform Distribution (pmf)

① Continuous Uniform Distribution { Continuous Random Variable }

Eg: The number of candies sold daily at a shop is uniformly distributed

[15-30] [Max, Min] \Rightarrow Interval



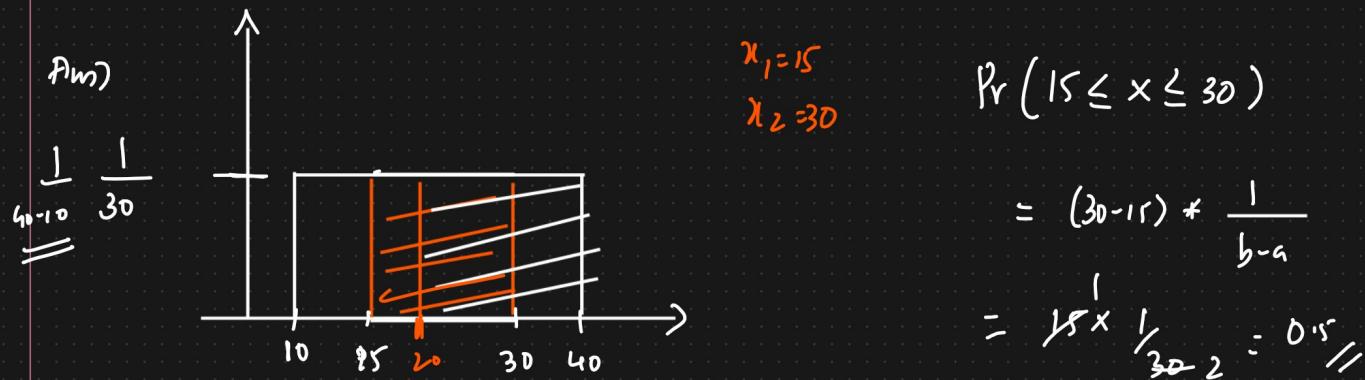
Notation: $U(a, b)$ $b > a$

Parameters: $-\infty < a < b < \infty$

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

Eg: The number of Landins sold daily at a shop is uniformly distributed with a maximum of 40 and a minimum of 10.

(i) Probability of daily sales to fall between 15 and 30.

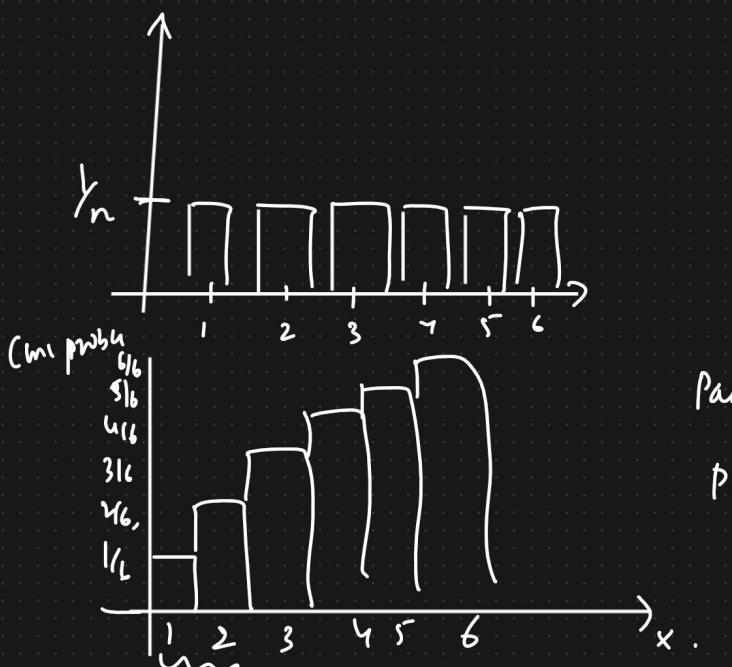


$$\Pr(x \geq 20) = (40 - 20) * \frac{1}{30} \\ = \frac{20}{30} = 66.66\%$$

② Discrete Uniform Distribution {Discrete Random Variables}.

Rolling a dice = {1, 2, 3, 4, 5, 6}

$$\Pr(1) = \frac{1}{6}, \quad \Pr(2) = \frac{1}{6}$$



$$n = b - a + 1$$

$$n = 6 - 1 + 1 = 6$$

Notation $U(a, b)$

Parameters a, b with $b > a$

PMF $\frac{1}{b-a}$

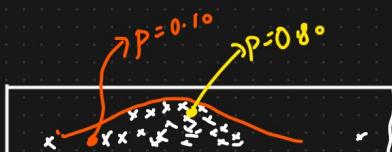
$$\Pr(X = 1 \text{ or } 2)$$

we will study F distribution/ chi square later in detail. For now, you should must search on internet and get some idea.

Hypothesis Testing

[Inferential Stats]

(1) P value



Out of 100 touches in this Space bar 10 times touching in that area.

Hypothesis Testing

Person \rightarrow Crime

① Null Hypothesis H_0 - Person has not committed Crime.

Alternate Hypothesis H_1 - Person has committed Crime

② Experiments : Proofs, DNA, fingerprints, evidence \Rightarrow Judge \Rightarrow Person has committed Crime

③ Reject the Null Hypothesis

OR

We fail to Reject H_0 .

① Coin is fair or No through 100 experiments

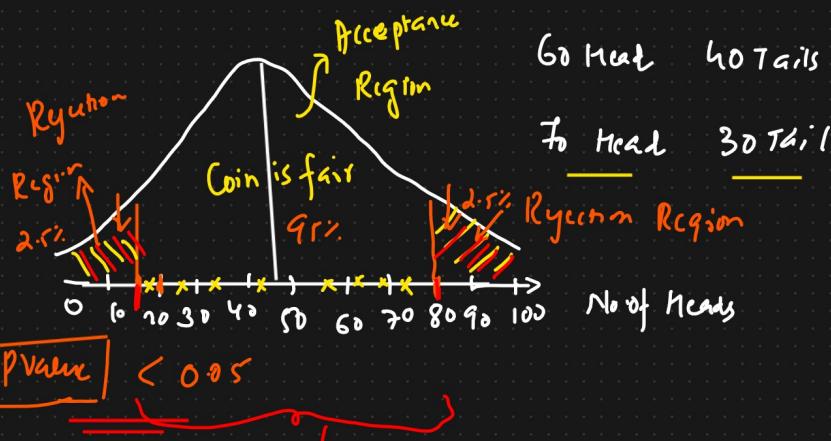
H_0 = Coin is fair

H_1 = Coin is not fair

Experiment : C.I is given by experts

confidence interval \rightarrow C.I = 95%

Significance level $\rightarrow \alpha = 1 - C.I = 0.05$



50 Head 50 Tails

60 Head 40 Tails

70 Head 30 Tails

Rejection Region

↓ ↓
probability Confidence Interval
value for
the Null Hypothesis
To be True

Assignment : Exploring Distribution