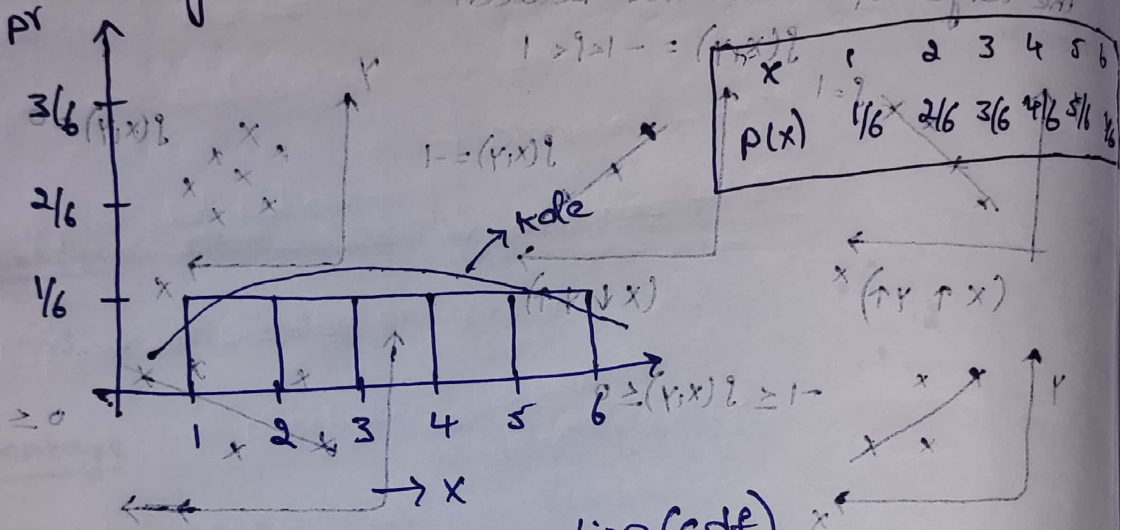
Probability distribution functions:-

Probability mass function:-

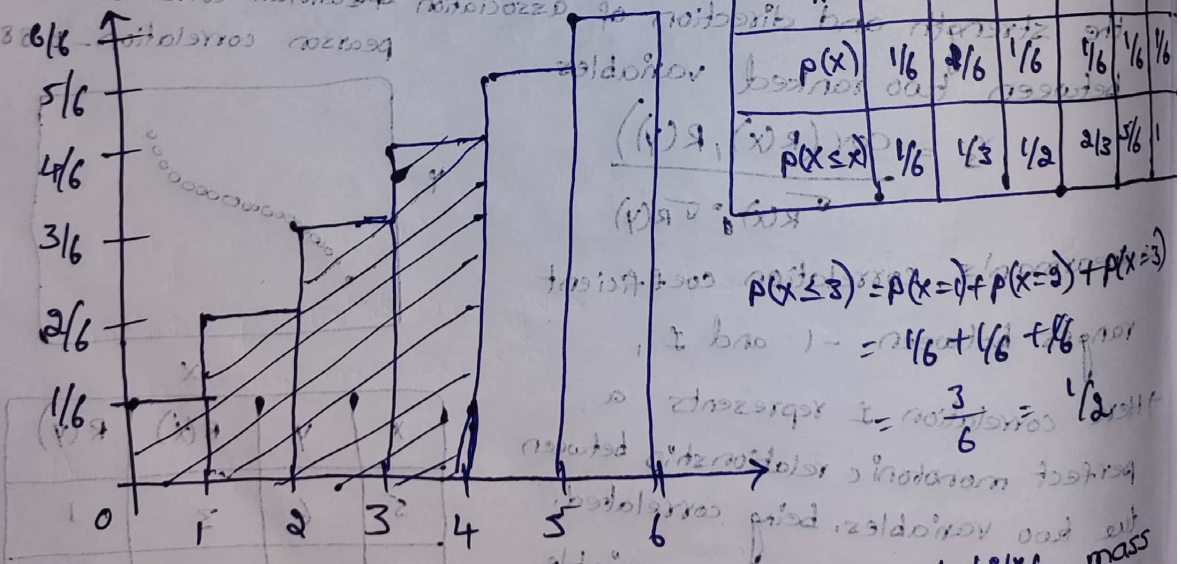
→ A probability mass function is a function that gives the probability that a discrete random variable is exactly equal to some value.

Ex:- Rolling a dice {1, 2, 3, 4, 5, 6}



x	1	2	3	4	5	6
p(x)	1/6	2/6	3/6	4/6	5/6	1/6

Cumulative distribution function (cdf)



x	1	2	3	4	5	6
p(x)	1/6	2/6	1/6	1/6	1/6	1/6
p(x≤x)	1/6	1/3	1/2	2/3	5/6	1

$P(x \le 3) = P(x=1) + P(x=2) + P(x=3)$

$= 1/6 + 1/6 + 1/6$

$= \frac{3}{6} = 1/2$

→ The cumulative distribution function (CDF) for a probability mass function (PMF) is function that describes the probability of a discrete random variable taking a value less than or equal to a specific value.

→ cdf at a given value of x is equal to sum of all possible outcomes up to and including a certain

2) probability density function:-

In probability theory, the probability density function (pdf) is a function that describes the probability distribution of a continuous random variable. It gives the probability that a random variable x takes on a value in a certain interval denoted by $p(a <= x <= b)$.

Fig:- pdf

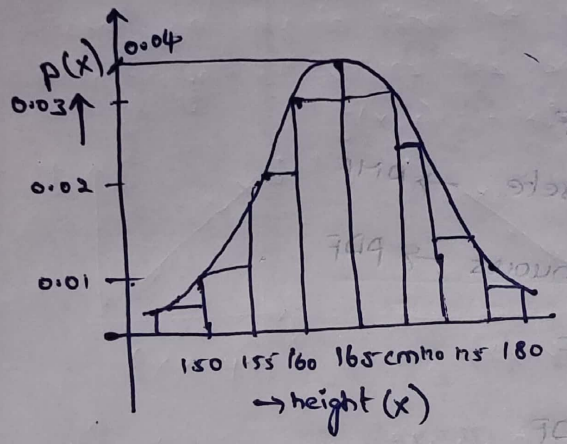
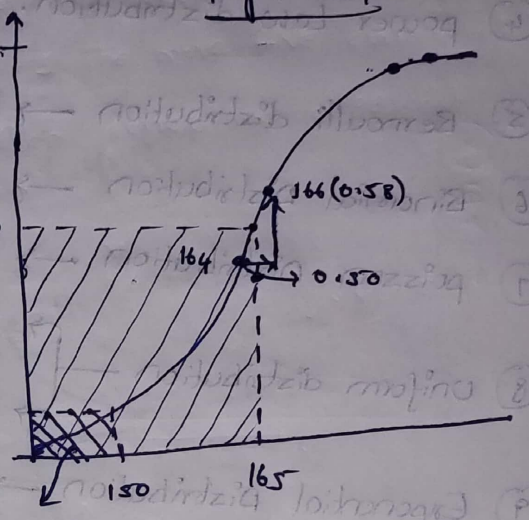


Fig:- CDF



$$cdf = \frac{0.58 - 0.50}{166 - 164} = \frac{0.08}{2} = 0.04$$

Relation between pdf and cdf:-

→ Here the cdf at any given point x is equal to the integral of the pdf from negative infinity upto x.

→ In other words, the cdf gives the probability that a random variable x is less than or equal to a certain value x.

→ while pdf gives the probability density of x at that value.

→ The pdf can be obtained from the cdf by differentiating the CDF w.r.t x:

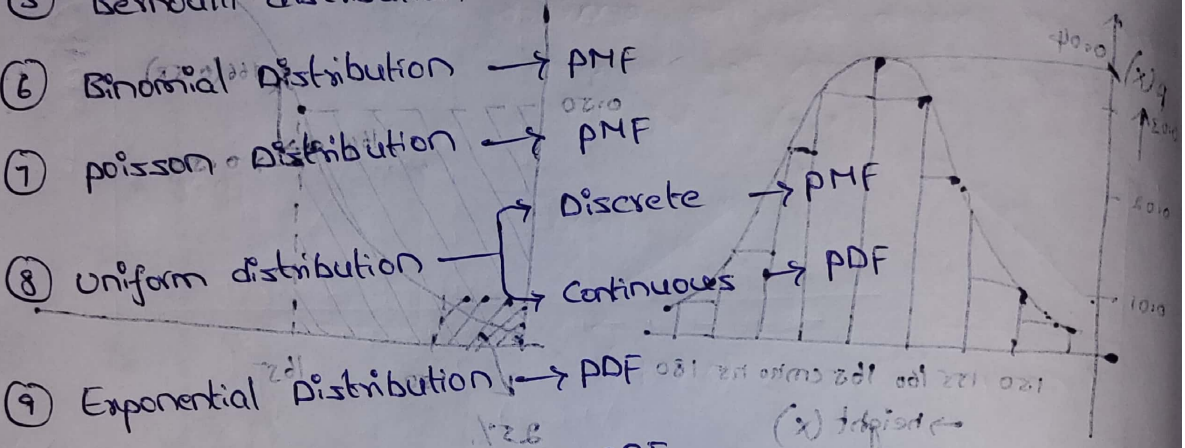
$$PDF(x) = \frac{d}{dx} CDF(x)$$

Ex:- The pdf is used find the probability of x falling within a certain range of values

→ The probability of students being present in 160 - 170 cm tall. [illegible] in a class of [illegible].

Probability Distributions

1. Normal / Gaussian Distribution → PDF
2. Standard Normal distribution → PDF
3. Log Normal distribution → PDF
4. power Law distribution →
5. Bernoulli distribution → PMF
6. Binomial Distribution → PMF
7. poisson Distribution → PMF
8. Uniform distribution → Discrete → PMF; Continuous → PDF
9. Exponential Distribution → PDF
10. CHI-square Distribution → PDF
11. F Distribution



probability distribution :-

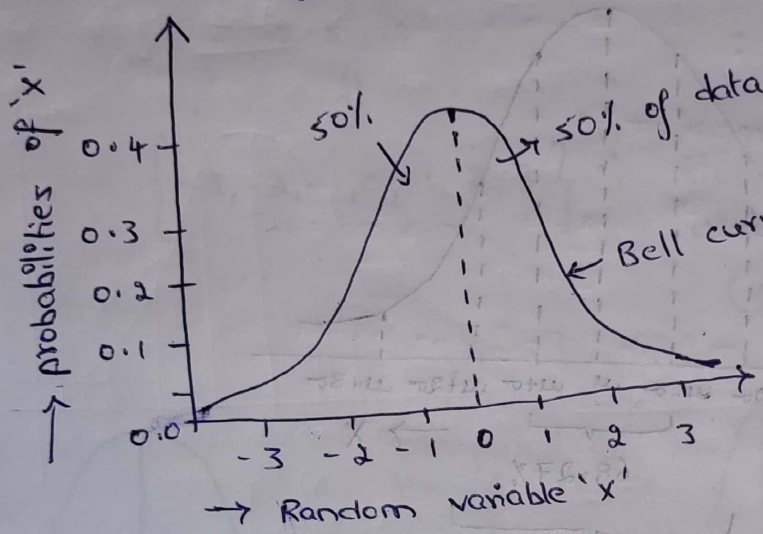
The probability distribution gives the possibility of each outcome of a random experiment or event.

① Normal/Gaussian Distribution:

Let X be the random variable follows Gaussian distribution with mean (μ) and standard deviation (σ)

Here, X is a continuous random variable.

Fig: Normal distribution.



properties of Normal distribution:

Following are properties of X which is following the Normal distribution:

→ In a normal distribution, the mean, median and mode are equal (i.e; Mean = Median = Mode).

→ The total area under the curve should be equal to 1.

→ The Normally distributed curve should be symmetric at the centre.

→ There should be exactly half of the values are to the right of the centre and exactly half of the values are to the left of the centre.

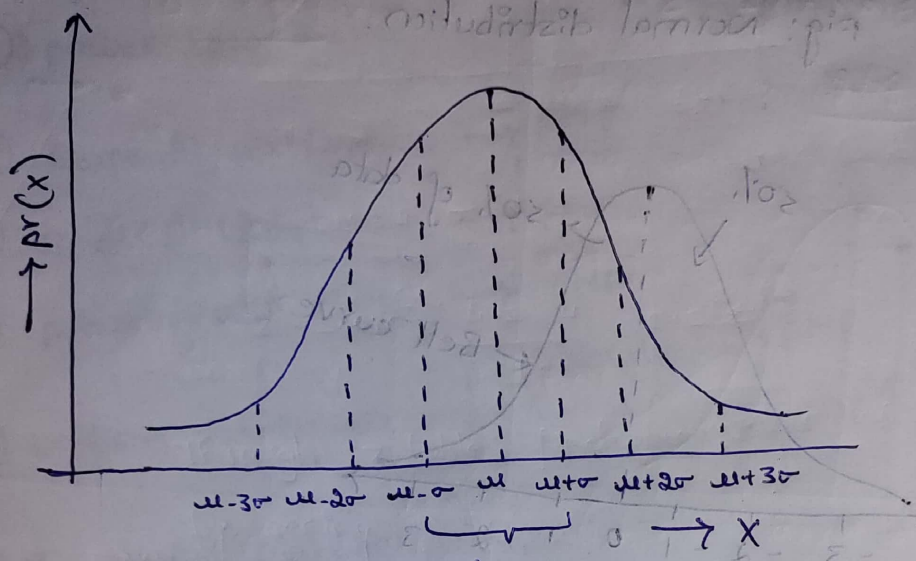
→ The normal distribution should be defined by the mean and standard deviation.

→ The Normal distribution curve must have only one peak.

→ The curve approaches the x-axis, but it never touches, and it extends farther away from the mean.

Empirical Rule:

The empirical rule is the statistical rule that almost all normal distributed data fall within three standard deviations of the mean.



$pr(\mu-\sigma \leq X \leq \mu+\sigma) \approx 68.27\%$ within $1^{st}\sigma$

$pr(\mu-2\sigma \leq X \leq \mu+2\sigma) \approx 95.45\%$ within $2^{nd}\sigma$

$pr(\mu-3\sigma \leq X \leq \mu+3\sigma) \approx 99.47\%$ within $3^{rd}\sigma$

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

Examples:

1. Height
2. Rolling a dice
3. Tossing a coin
4. IQ
5. Blood pressure
6. shoe size
7. Weight

2) Standard Normal distribution :

A Normal distribution with zero mean ($\mu = 0$), unit variance ($\sigma = 1$) is called as standard Normal Distribution.

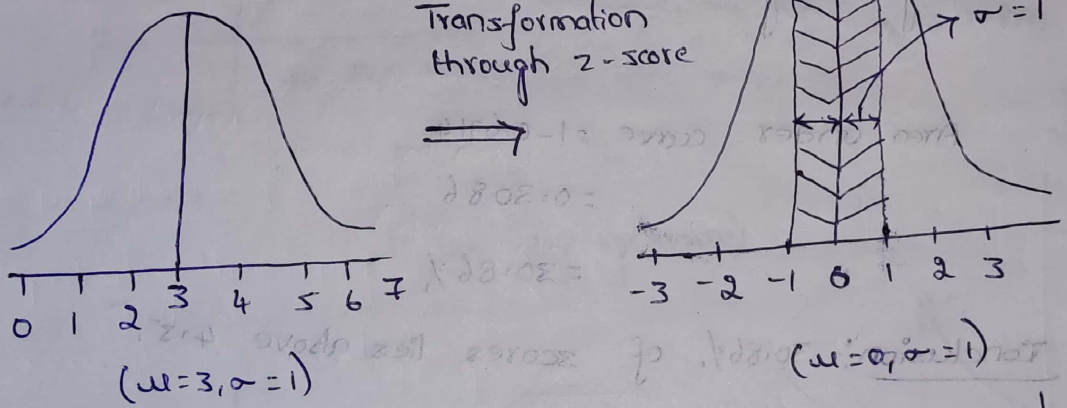
$\{0, 1, 2, 3, 4, 5, 6, 7\} \Rightarrow$ Normal distribution. $(\mu = 3, \sigma = 1)$

$\Downarrow$

z-score

$\Downarrow$

$\{-3, -2, -1, 0, 1, 2, 3\} \Rightarrow$ standard Normal distribution. $(\mu = 0, \sigma = 1)$



$(\mu = 3, \sigma = 1)$ Normal distribution

$(\mu = 0, \sigma = 1)$ standard Normal distribution.

z-score :

z-score tells you about a value how many standard deviation away from the mean.

$$z_{score} = \frac{x_i - \mu}{\sigma}$$

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

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

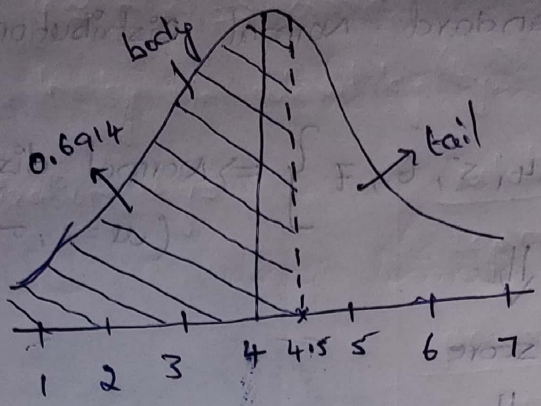
$\frac{3-3}{1} = 0$

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

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

$\frac{6-3}{1} = 1$

Ex:- What is the percentage of scores which lies above 4.5?



$\mu = 4, \sigma = 1$

$$z\text{-score} = \frac{4.5 - 4}{1} = 0.5$$

Through z-table find the value of z-score by seeing in it.

Area under curve $= 1 - 0.6914$

$= 0.3086$

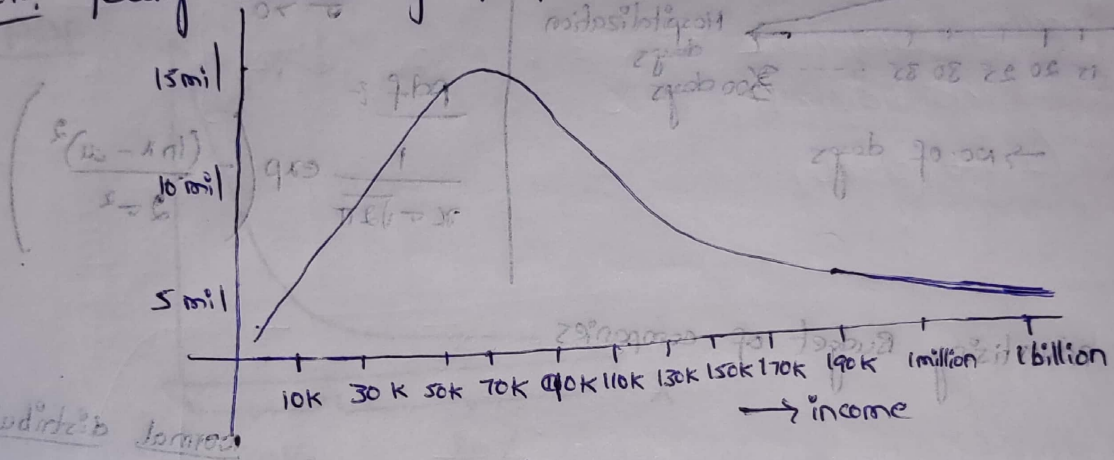
$= 30.86\%$

Conclusion:- 30.86% of scores lies above 4.5.

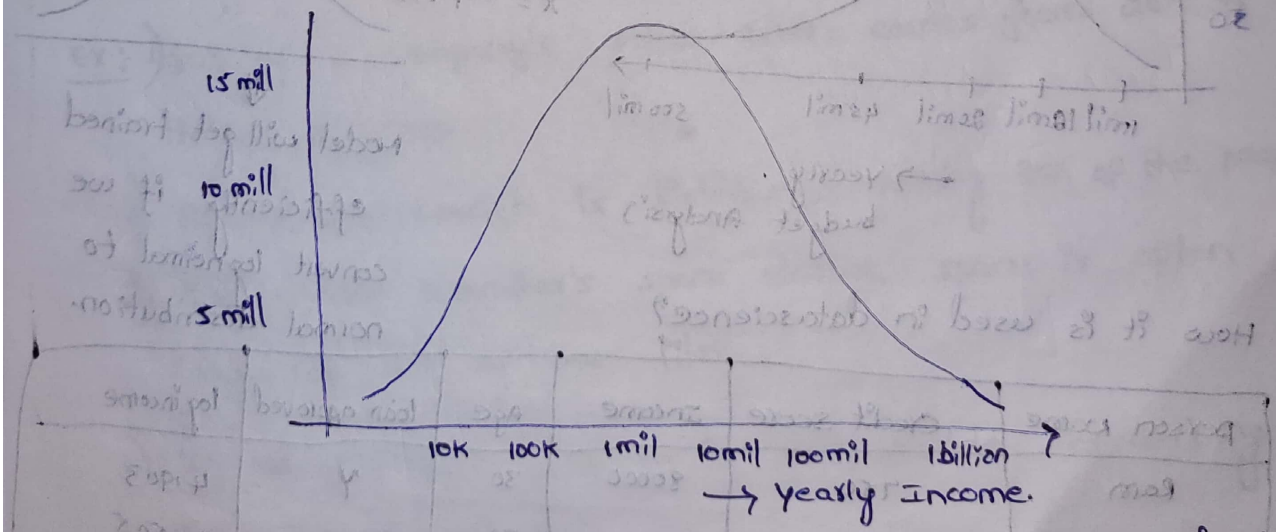
(3) Log Normal Distribution:

In probability theory, a log-normal 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, then the exponential function of $Y, X = \exp(Y)$, has a log-normal distribution.

Ex:- yearly income of people.



if we do log(income)



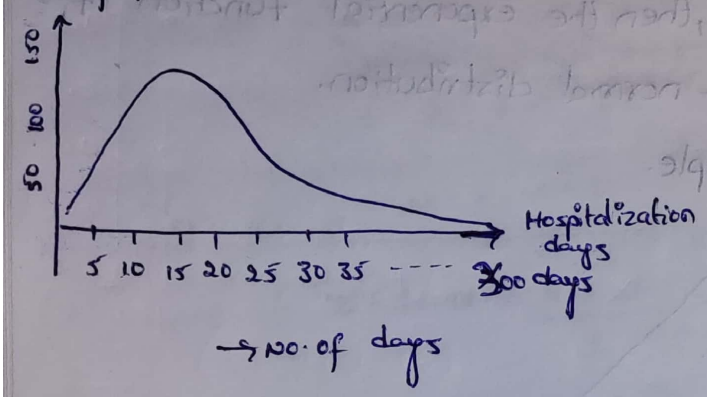
→ If you get a normal distribution by applying a log function to a dataset then dataset is log normally distributed.

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

$y \approx \ln(x) \Rightarrow$ Normal distribution(μ, σ^2)

$\log \approx \log_{10}$ $\ln$ = natural $\log (\log_e)$

Ex 2: Hospitalization days



notation

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

parameters

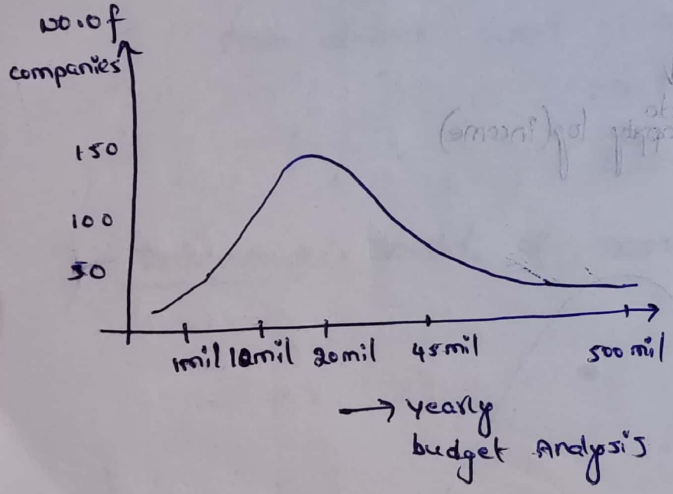
$\mu \in (-\infty, +\infty)$

$\sigma > 0$

pdf:

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

Ex 3: Advertising Budget of companies

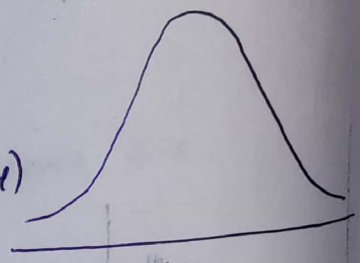


Normal distribution

logarithm transformation

$\ln(x)$ →

← $x = \exp(y)$



⇓

Model will get trained efficiently if we convert lognormal to normal distribution.

How it is used in datascience?

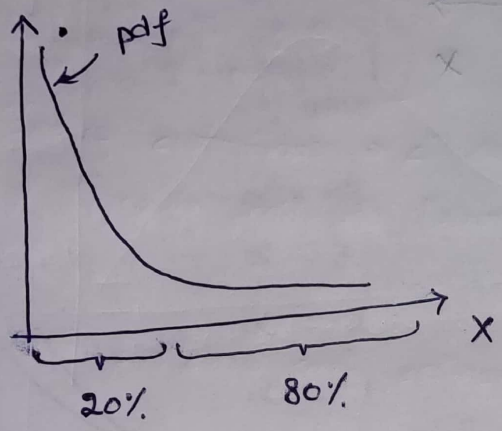
person name	Credit score	income	Age	loan approved	log income
Ram	750	80000	30	Y	4.903
Sam	310	32000	32	N	4.505
Tom	475	77000	40	Y	4.886
Dinesh	600	65000	35	N	4.812
Hari	820	550000	37	Y	5.740
Varun	780	75000	22	Y	4.8750

quite different

④ power law distribution :-

In statistics, power law is a functional relationship between two quantities, where a relative change in one quantity results in a proportional relative change in other quantity, independent of initial size of those quantities one quantity varies as power of another.

PDF



→ A power law distribution describes the ~~80-20~~ 80-20% rule that governs many phenomena around us.

Ex :- 1) 80% of a company's sales often comes from 20% of their customers.

2) 80% of the wealth is distributed among 20% of the people.

3) 80% of the computer's ~~space~~ storage space is often taken by 20% of the files.

4) In ipl, 20% of the team is responsible for winning 80% of match.

5) 80% of project complete by 20% of team.

Types of power law distribution :-

① pareto Distribution ② Exponential distribution.

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.

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

PDF

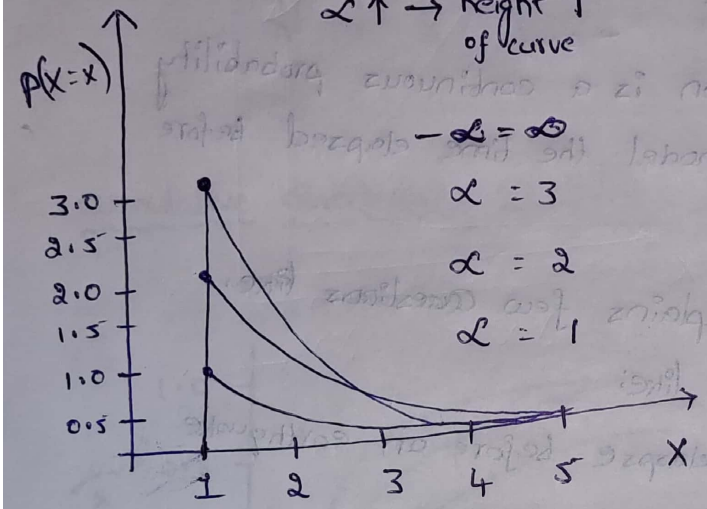
$\alpha \uparrow \rightarrow$ height $\uparrow$ of curve

$-\alpha = \infty$

$\alpha = 3$

$\alpha = 2$

$\alpha = 1$



$x_m = 1$

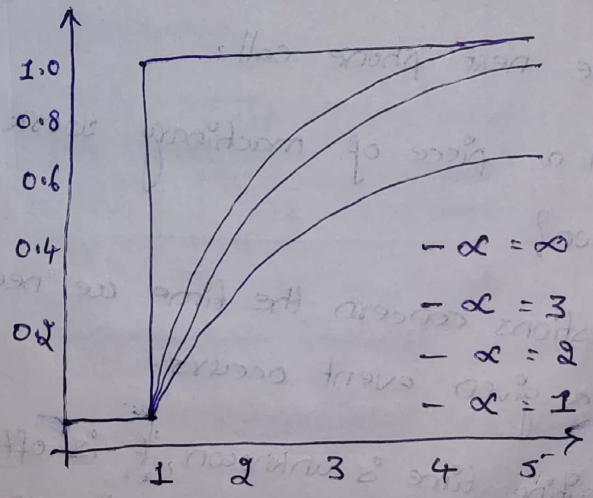
$pr(x \geq 4) = \left(\frac{1}{4}\right)^2 = \frac{1}{16}$

$pr(x \geq 2) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$

$= 0.25$

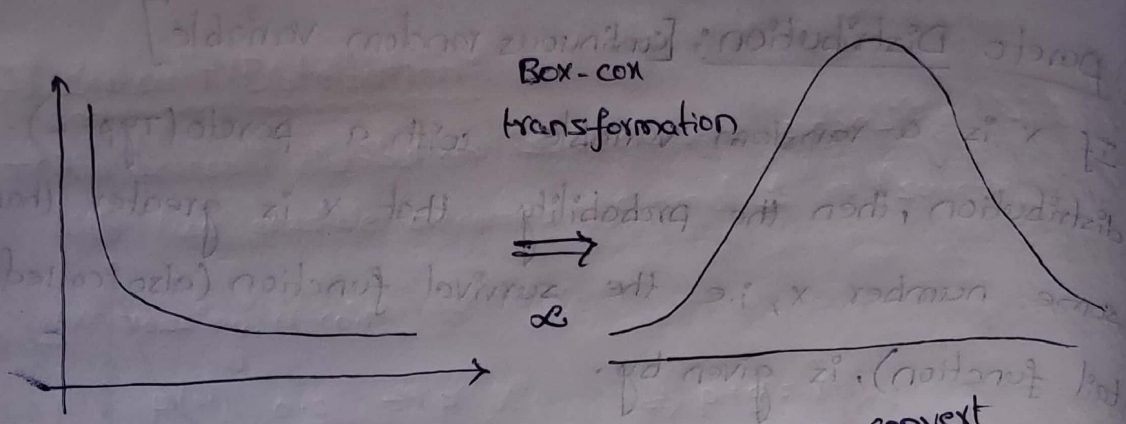
$x_m = 2$, then height will increase more.

→ pareto type I probability density functions for various α with $x_m = 1$. As $\alpha \to \infty$, the distribution approaches $\delta(x - x_m)$ wher δ, is dirac delta function.



pareto Type I with cumulative distribution functions

→ pareto distribution follows power law (80-20)% rule.



By applying Box-cox transformation we can convert pareto distribution to normal distribution.

(2) Exponential Distribution:-

The exponential distribution is a continuous probability distribution used to model the time elapsed before a given event occurs.

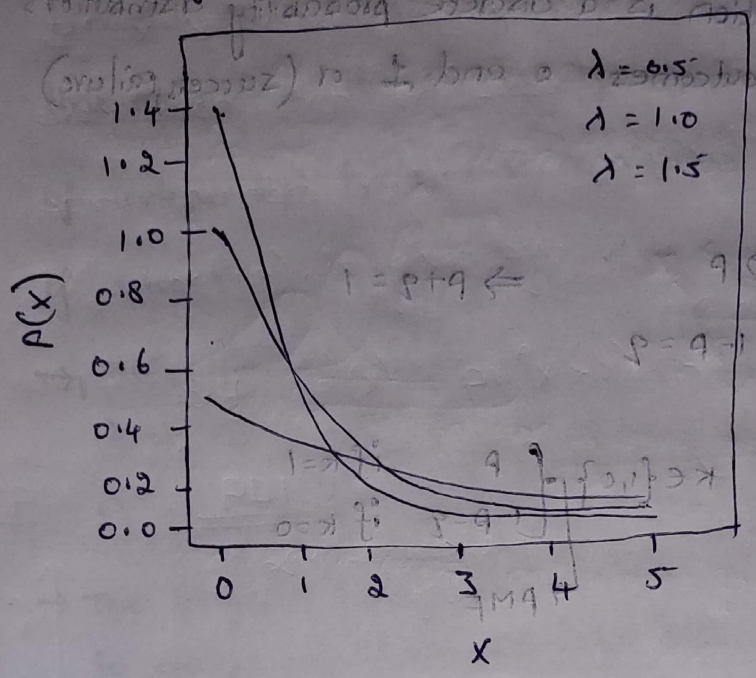
→ This distribution explains few ~~Questions~~ like:

1) How much time will elapse before an earthquake occurs in a given region.
2) How long do we need to wait until a customer enters our shop?
3) How long will it take before a call center receives the next phone call.
4) How long will a piece of machinary work without breaking down?

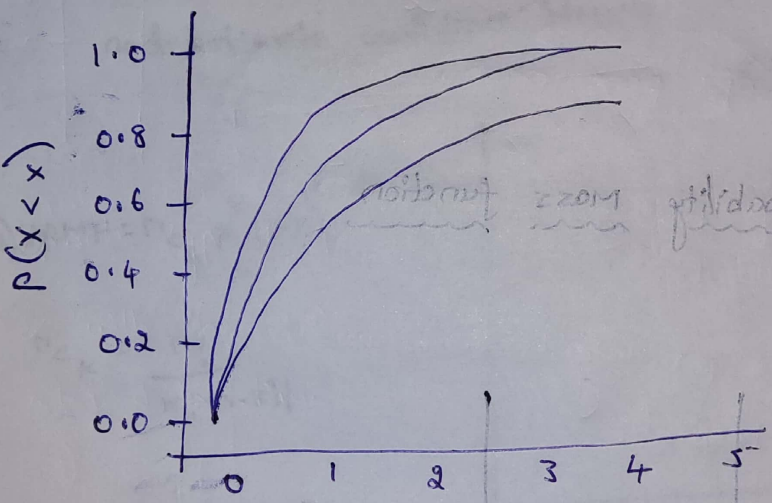
All these questions concern the time we need to wait before a given event occurs.

→ If this waiting time is unknown, it is often appropriate to think of it as a random variable having an exponential distribution.

probability Density function:



Cumulative distribution Function:



$$PDF = \lambda e^{-\lambda x}$$

$$CDF = 1 - e^{-\lambda x}$$

(5) Bernoulli Distribution:-

→ The bernoulli distribution is a discrete probability distribution with two possible outcomes 0 and 1 or (success, failure)

Ex:- Tossing a coin

$pr(H) = 0.5 \Rightarrow p$ $\Rightarrow p+q = 1$

$pr(T) = 0.5 \Rightarrow 1-p = q$

$$PMF = p^k(1-p)^{1-k}$$

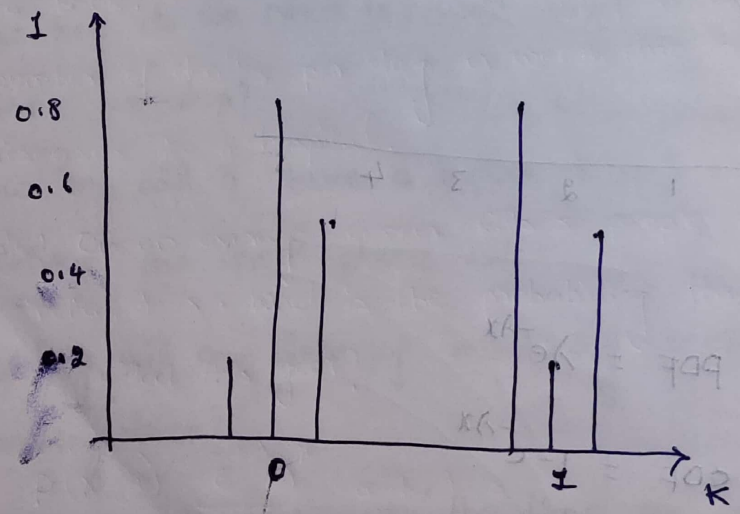
$k=1 \Rightarrow p^1(1-p)^0 = p$

$k=0 \Rightarrow p^0(1-p)^{1-0} = 1\times(1-p) = 1-p$

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

↳ PMF

probability Mass function



Three examples of bernoulli distribution.

$P(x=0) = 0.2$ and $P(x=1) = 0.8$

$P(x=0) = 0.8$ and $P(x=1) = 0.2$

$P(x=0) = 0.5$ and $P(x=1) = 0.5$

(6) Binomial distribution:

Binomial distribution is a discrete probability distribution that describes the number of successes in a fixed number of independent trials, where each trial has the same probability of success.

→ This distribution is characterized by two parametres:

1) The probability of success in each trial (denoted by p)
2) The number of trails (denoted by n).

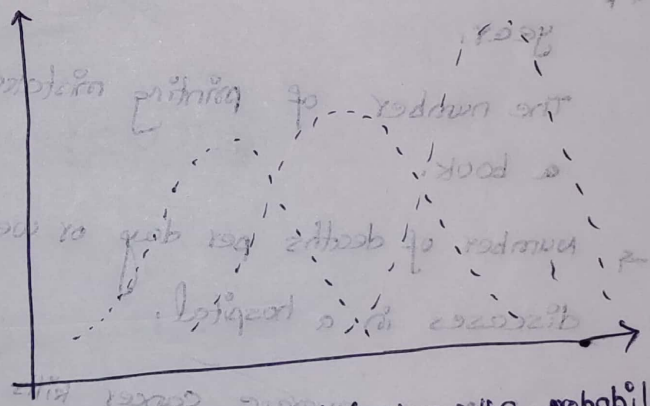
→ The binomial distribution is often used in data science to model outcomes that have two possible outcomes, such as success or failure, yes or no, or heads or tails.

→ It is used to calculate probabilities of various outcomes and can be used to perform hypothesis testing, make predictions and estimate confidence intervals.

PMF

$$P(X=k) \text{ or PMF} = {}^{n}C_{k}\, p^{k}(1-p)^{n-k}$$

$${}^{n}C_{k} = \frac{n!}{k!(n-k)!}$$



Ex: 1 suppose we flip a fair coin 10 times and want to know the probability of getting exactly 3 heads. Here n=10 and p=0.5

A) X represents the number of heads obtained in 10 coin flips

k = 5

n = 10

p = 0.5 since the coin is fir

$${}^{n}C_{k} = \frac{10!}{3!7!} = \frac{10\times 9\times 8\times \cancel{7\times 6\times 5\times 4\times 3\times 2\times 1}}{3\times 2\times 1\ (\cancel{7\times 6\times 5\times 4\times 3\times 2\times 1})}$$

$$= \cancel{8}$$

$$P(X=3) = {}^{10}C_{3} \times 0.5^{3} * 0.5^{(10-3)}$$

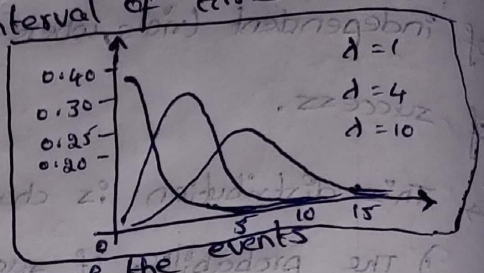
$$= \cancel{2000} = 120 \times 0.5^{10}$$

$$\approx 0.117$$

probability of getting exactly 3 heads in 10 flips is 11.7%

7) poisson distribution :- (discrete random variable)

→ The poisson distribution is a probability distribution that models the number of times an event occurs within a specific interval of time or space



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

$e = 2.7182$

λ - Is the average rate of the events per interval.

x - no. of events that occur in a given interval.

→ poisson distribution is used in cases where the chances of any individual event being success is very small.

Ex:- The number of plane crash in india in one year.

The number of printing mistakes in each page of a book.

→ Number of deaths per day or week due to rare diseases in a hospital.

Example:- On an average cancer kills 5 people each year in the india, $\lambda = 5$. What is the probability that only one person is killed this year?

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

$$P(x = 1) = \frac{e^{-5} 5^{1}}{1!} \approx 0.033$$

Ex:- The number of people visits bank in an hour.

(8) Uniform distribution:

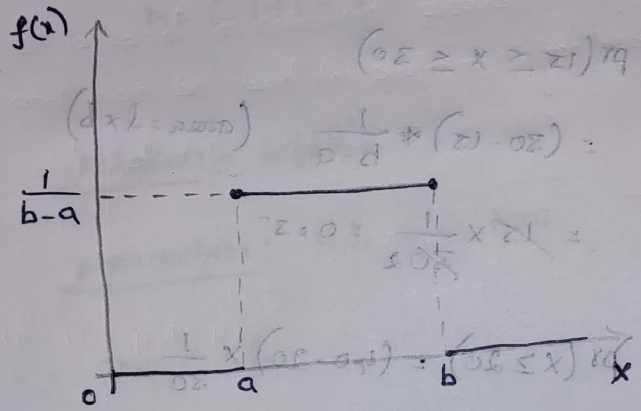
(1) continuous uniform distribution. (pdf)

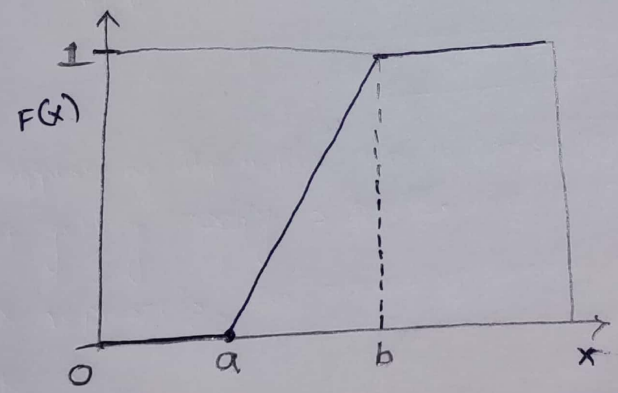
(2) ~~Uniform~~ Discrete uniform distribution. (pmf)

i) Continuous uniform distribution:-

A continuous uniform distribution is a probability that is used to describe situations where all values within a certain range are equally likely to occur.

→ In other words, the probability density function of a continuous uniform distribution is a constant within a specified interval and zero outside that interval.





→ The probability density function (pdf) of a continuous uniform distribution with parametres a and b is given by:

Notation:- [Max, Min] ⇒ Interval

$U(a,b)$

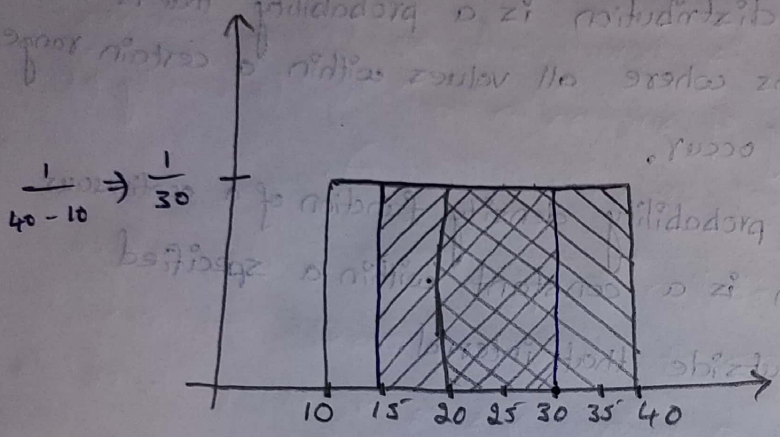
parametres:- $-\infty < a < b < \infty$

$b >> a$

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

Eg:- The no. of candidates sold daily at a shop is uniformly distributed with a maximum of 40 and minimum of 10.

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



$x_1 = 15$

$x_2 = 30$

$pr(15 \le x \le 30)$

$= (30-15) * \frac{1}{b-a}$ (area = l × b)

$= 15 \times \frac{1}{30} = 0.5$

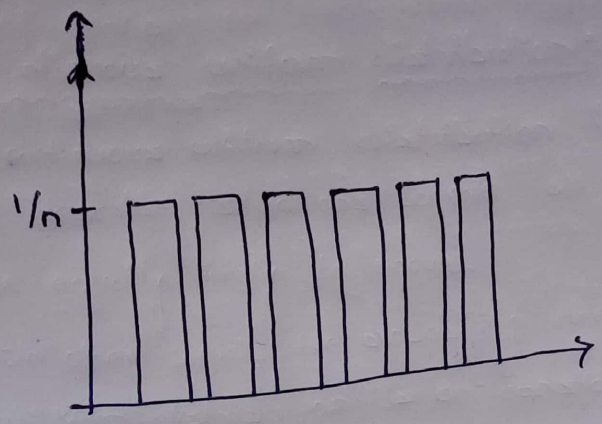
$pr(x \ge 20) = (40-20) \times \frac{1}{30}$

$= 20 \times \frac{1}{30} = 66.66\%$

ii) Discrete uniform distribution :-

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

pr(1) = 1/6 pr(n) = 1/6



$n = b - a + 1$

$n = 6 - 1 + 1 = 6$

Notation :- $u(a, b)$

parametres :- a, b with $b > a$

PMF : $1/n$