Probability dist- ⇒ ① PMS | CDF Probability Dist- With Discreate R.V.

✓ ② PDF/CDF Probability dist with Continues R.V.

③ CDF

PMS

① Uniform (Coin, Dice)

② Bernoli Dist (Coin) ⇒ Binary outcome

③ Binomial Dist- ⇒ n try with binary outcome

④ multinomial Dist- ⇒ n try with multiple outcome

✓ ⑤ Possion Distribution (time)

⇒ CDF (Commulative Dist- function)

PDF ⇒

① Uniform Dist-

② normal Distribution

③ SND = ⇒ imp

④ Log normal Dist-

⑤ exponantial Dist-

⑥ Powerlaw Dist-

⇒ Skewness

⇒ CDF

Power law

Pareto | exponantial

(7) t-distr, F-distr, Chi-Square Distr

Probability Distribution.

(1) Probability Density function

$$F(x) = P(a \le x \le b) = \int_a^b f(x)\,dx \ge 0$$

Range

$[a, b]$

(2) Probability Mass function

$$P(x) = P(X = x)$$

Random var. (Discreate Random var)

(3) Cumulative Distribution function (CDF)

$$F(x) = P(X \leq x)$$

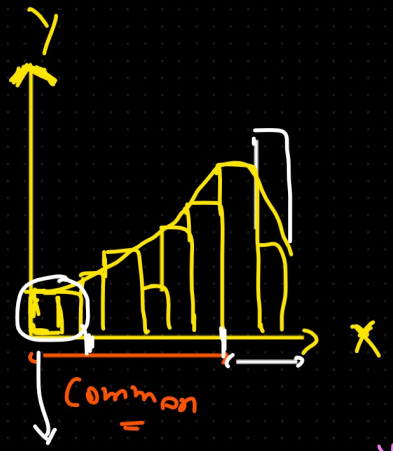
H H ⇒
H T ⇒
T H ⇒
T T ⇒

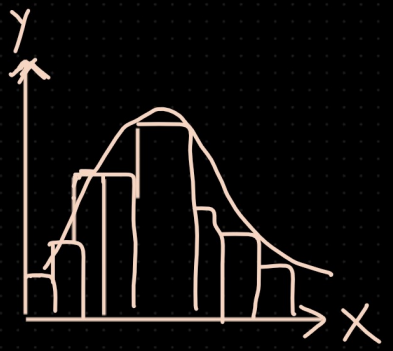
$0 \leq x \leq 2$

$H \leq 2$

$H = 0$
$H = 1$
$H = 2$

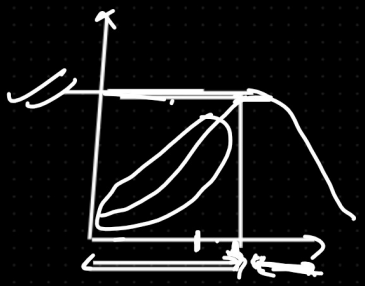
Data ⇒ Distributed (spread out) in different way.

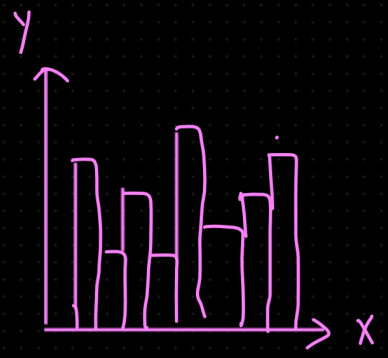


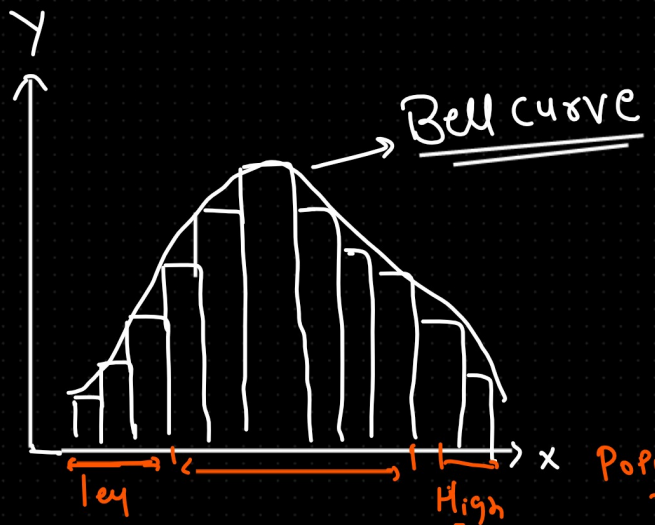


left skwed Data

1. Income of the people
2. retirment age of people







Population =

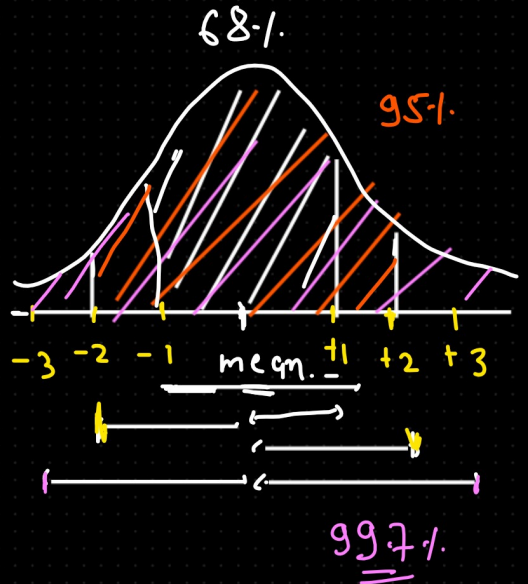
1. Height of People
2. Size of things Produce by machine.
3. Errors in measurement
4. blood Pressure
5. Marks in the test

mean = median = mode

(Perfect normal Distr)

Symmetry

Empirical rule



$\Rightarrow$ 100%.

(1) S.D.

if my Data is normally Distr

68%.

$= \boxed{\text{Mean} \pm 1\sigma}$

$\Downarrow$

Spread of the Data on this X-axis

(2) S.D. $\Rightarrow$ 95%.

$\boxed{\text{mean} \pm 2\sigma}$

$\Downarrow$

Spread of the data on Y-axis

(3) S.D. $\Rightarrow$ 99.7%

$\boxed{\text{mean} \pm 3\sigma}$

$\Downarrow$

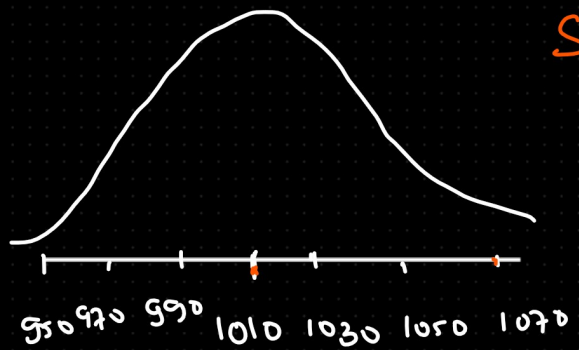
Spread of the on x-axis

Standard Score :- the number of S.D. from the mean is called Standard Score or Z-Score.

[950, 970, 990, 1010, 1030, 1050, 1070]

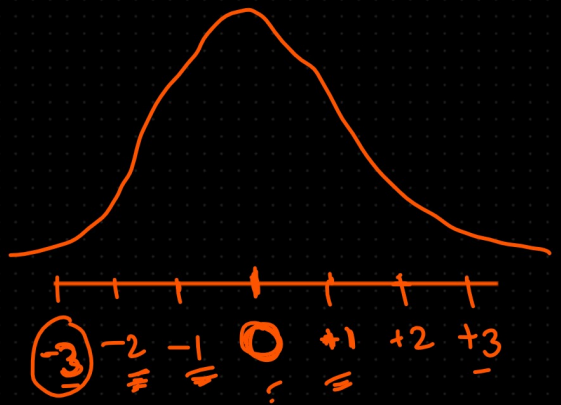
$\frac{X - \text{mean}}{S.D.}$

[mean = 0, Std dev = 1]



Normal Distr.

Standardization ⇒



Standard normal Distr.

max num ⇒ 60, min_threshold ⇒ 30

Class ⇒ [20, 15, 26, 32, 18, 28, 35, 14, 26, 22, 17]

= fail?

Range of the is high

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

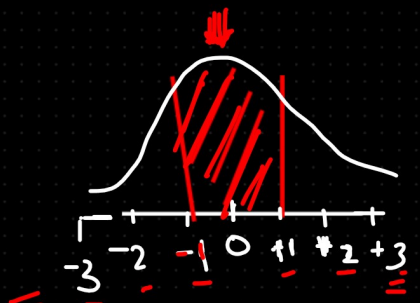
S.D.

$$\left[\frac{X-\mu}{\sigma}\right]$$

mean = 23

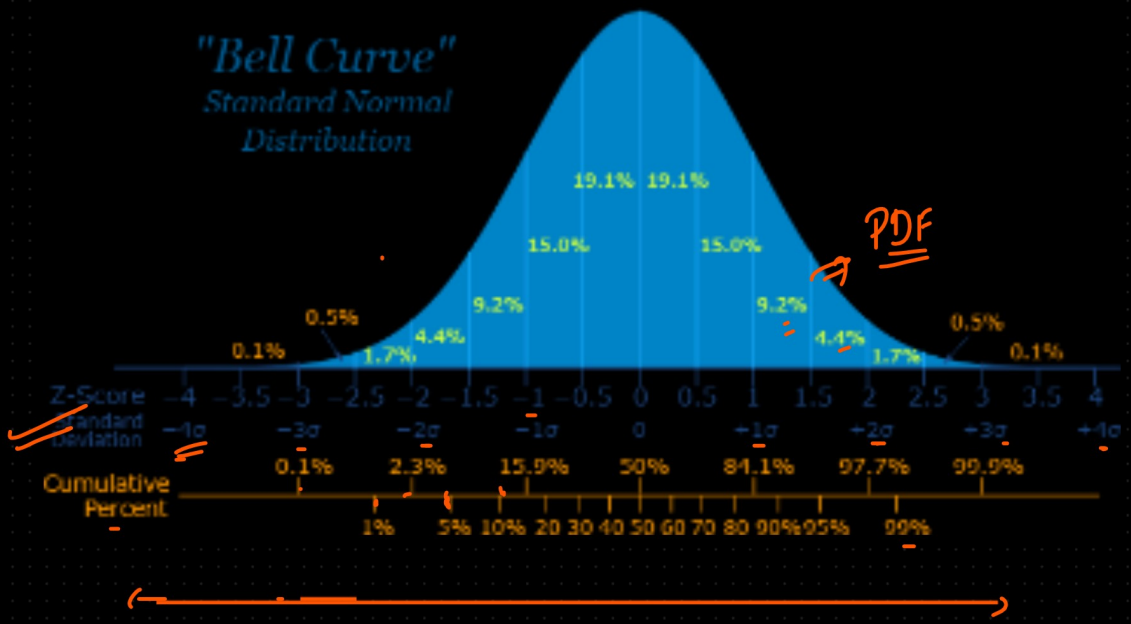
Std.dev = 6.6

-0.45, -1.21, 0.45, 1.36, -0.76, 0.76, 1.82, -1.36, 0.45, -0.15, -0.91

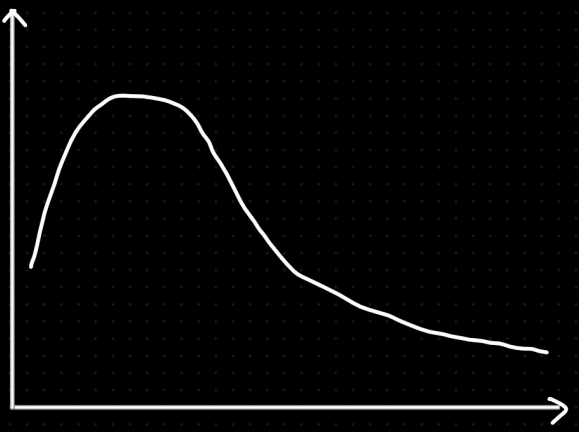


⇒ 68.1

1. Reduce the range
2. to interpretion of the Data wrt S.D.
3. zero centric


"Bell Curve"
Standard Normal Distribution
19.1%
19.1%
15.0%
15.0%
9.2%
9.2%
4.4%
4.4%
1.7%
1.7%
0.5%
0.5%
0.1%
0.1%
PDF
Z-Score
−4 −3.5 −3 −2.5 −2 −1.5 −1 −0.5 0 0.5 1 1.5 2 2.5 3 3.5 4
Standard Deviation
−4σ −3σ −2σ −1σ 0 +1σ +2σ +3σ +4σ
Cumulative Percent
0.1% 2.3% 15.9% 50% 84.1% 97.7% 99.9%
1% 5% 10% 20 30 40 50 60 70 80 90% 95% 99%

Log normal Distribution



right or pos Skewed
← Skewed data

Random var ⇒ Probability curve

⇓

log(x)

⇓

normal Distribution

Random variable

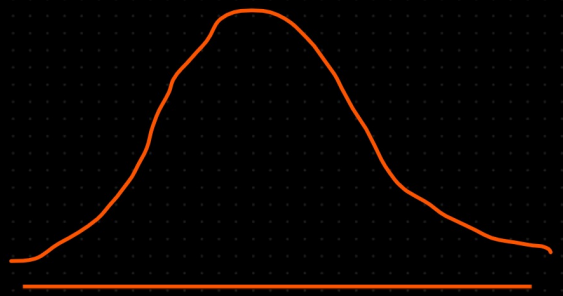
↑

log(R.V.)

↓

normal Dist.

Natural log ⇒ log(x) ⇒ e (exponential)



log → normal → exponential → log

$\Rightarrow X \sim \text{lognormal}(\mu, \sigma) \Rightarrow$ Normal Dist. (μ, σ)

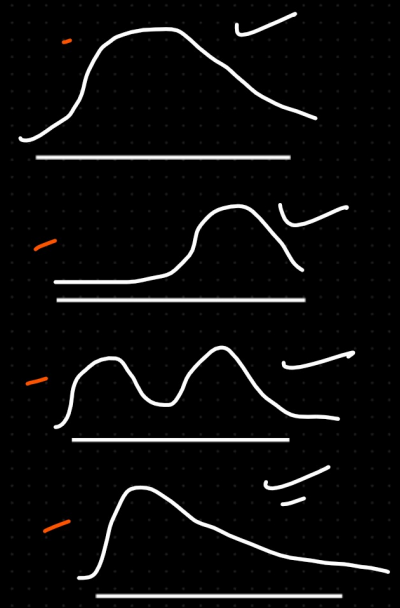
(ln = natural log (log e))

$Y = \ln(x)$

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.[1][2] Equivalently, if Y has a normal distribution, then the exponential function of Y, $X = \exp(Y)$, has a log-normal distribution. A random variable which is log-normally distributed takes only positive real values. It is a convenient and useful model for measurements in exact and engineering sciences, as well as medicine, economics and other topics (e.g., energies, concentrations, lengths, prices of financial instruments, and other metrics).

1. Power law
2. Pareto

} ⇒ 80% - 20% Rule



Power law Distr ⇒ Continues Random variable

(2-quantity)

One quantity → Second quantity

relative change

80/20

Ex:- ① 20% of the team is responsible for winning 80% of the matches.

② 20% Product drived 80% sale in the amazon.

③ 80% wealth is Distributed among 20% of people

④ 80% of Project completed by 20% of team.

⑤ 80% of tax Paid by 20% of Population.

Ex: of Power law Dist / type of the Power law ⇒ Pareto