[Day-3 Stats]

Page No.:
Date. / /

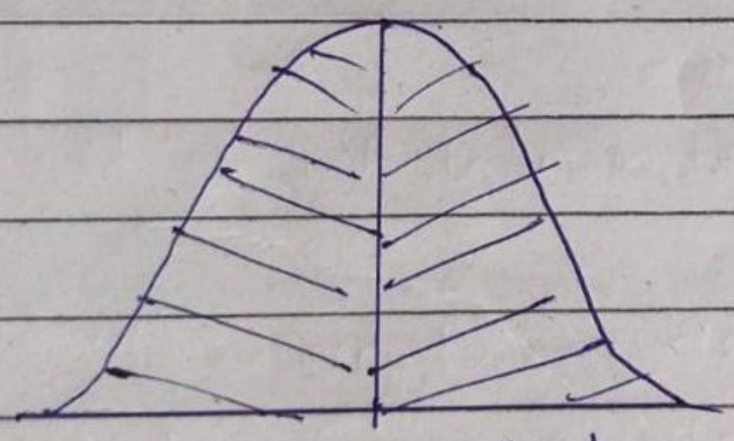
Topics

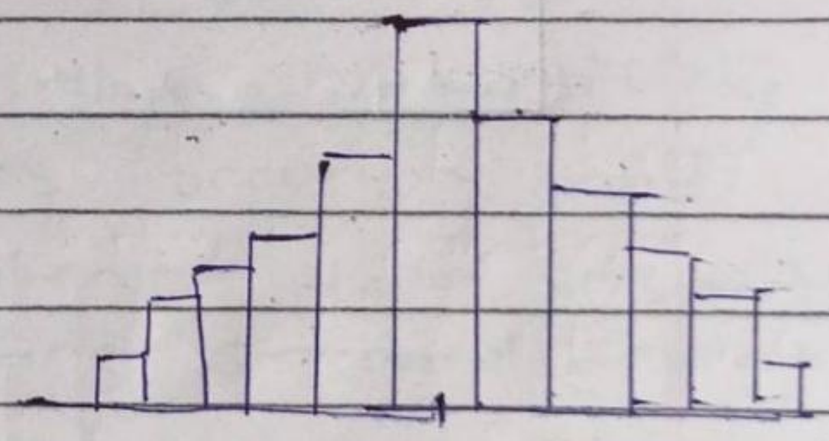
* Normal distribution
* Standard Normal distribution
* Z-score.

① Normal distribution :-

- also called as Gaussian distribution
- it is continous probability distribution that is symmetrical around its mean
- It describes how the values of variable are distributed.

example :- Ages, weight, height.





using kernal density estimator we can smooth histogram to normal distribution

Page No.:
Date: / /

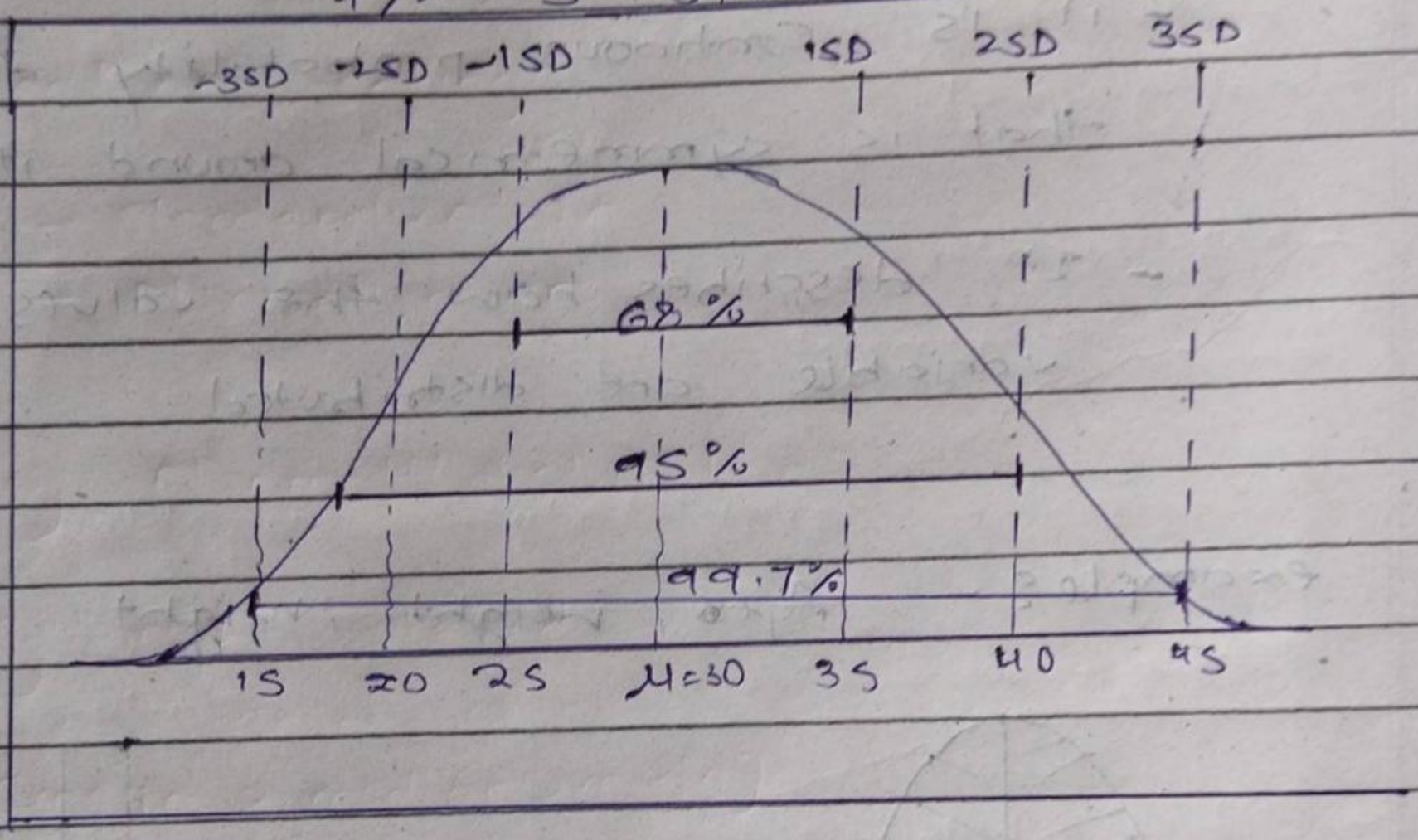
(*) empirical rule for normal distribution :-

empirical rule describe that in normal distribution :-

(1) 68% of the data fall within +/- 1 standard deviation.

(2) 95% of data fall within +/- 2 standard deviation

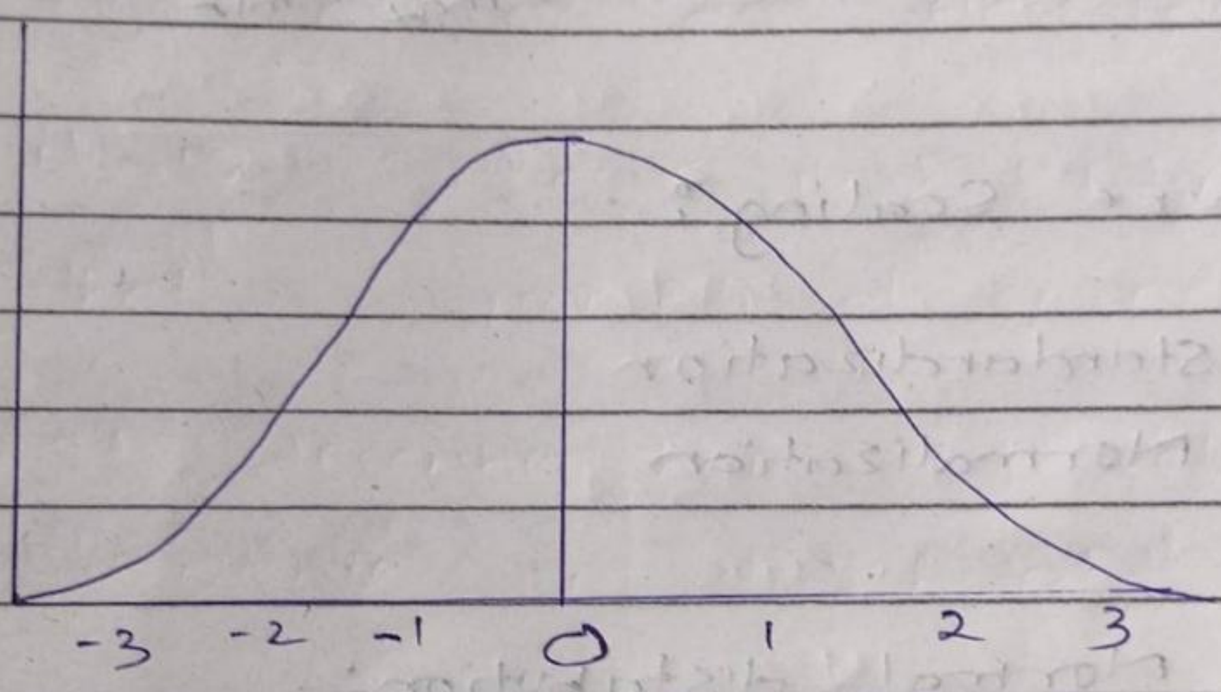
(3) 99.7% of data fall within +/- 3 standard deviation.



Page No.:
Date. / /

* Standard Normal distribution :-

- also called as z- distribution is a special normal distribution where the mean is 0 ($\mu = 0$) and standard deviation is 1 ($SD = 1$)



- converting normal distribution to standard normal distribution allows you to calculate the probability of certain values occuring and to compare different data sets with different mean and standard deviation

→ to convert we use z-score

$$z = \frac{x - \mu}{\sigma}$$

x = individual value

μ = mean

σ = standard daviation

Standardization :- All the data bring into same scale to reduce the complexity , $\mu = 0 \quad \sigma = 1$

Normalization :- [lower - scale <-> higher scale]

① min max Scalar [0-1]

$$\text{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

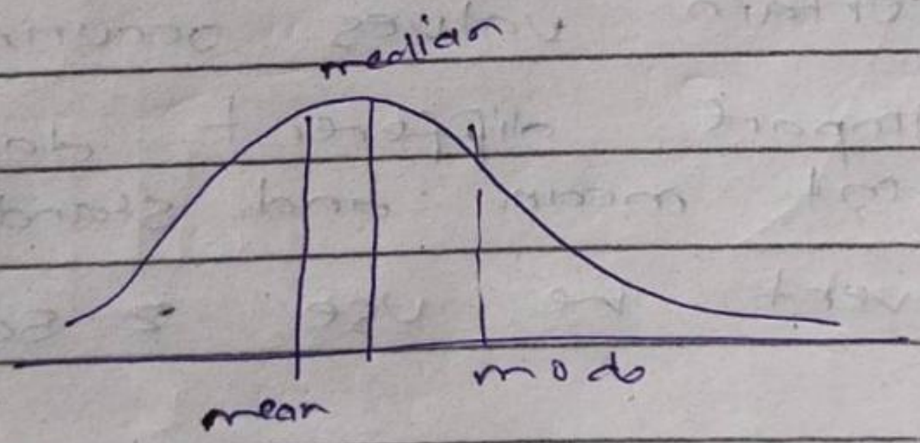
feature Scaling :-

- Standardization
- Normalization

log Normal distribution :-

log right skewed :-

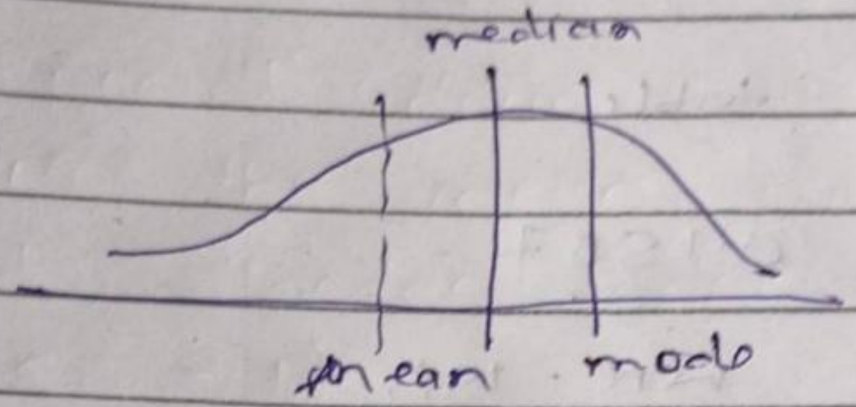
[mean > median > mode]



Page No.:
Date. / /

log left skewed :-

[mode > median > mean]



$x \sim$ log normal distribution

$y = \ln(x)$ $\{y = \log_e\}$ $\{\ln = \log_e\}$

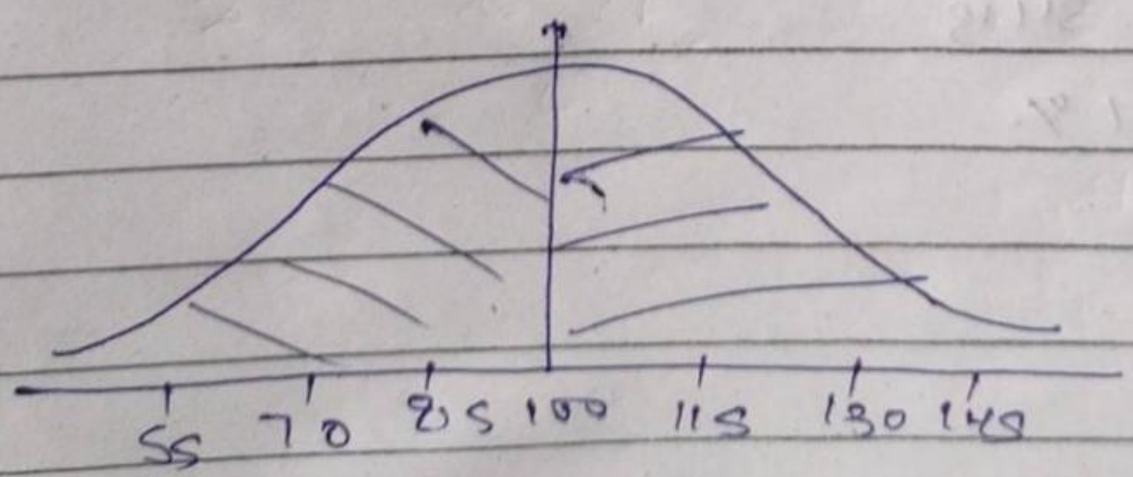
$= x \sim N(\mu, \sigma)$

$\Rightarrow y = \exp(x)$ { y is normal distribution}

In india the avg IQ is 100 with a SD of 15. what is the percentage of population would you expect to have an IQ.

(1) lower than 85
(2) higher than 85
(3) Between 85 and 100

=>



Page No.:
Date. / /

(1) lower 85 ÷

$$\Rightarrow \frac{x_i - \mu}{\sigma} = \frac{85 - 100}{15} = -1.00$$

By table

= 0.1587

here 15%.

(2) higher than 85 ÷

$$\Rightarrow \text{score} \Rightarrow \frac{x_i - \mu}{6} = \frac{85 - 100}{15} = -1.00$$

= By table

= 0.1587

= 1 - 0.1587

= 0.8413

= 84%.

(3) Between 85 and 100

By score for 85 ⇒ 0.1587

$$= \frac{x_i - \mu}{\sigma} = \frac{100 - 100}{15} = 0.00 = 0.5000$$

According to table

= 0.8413 − 0.5000

= 0.3413

= 34%.