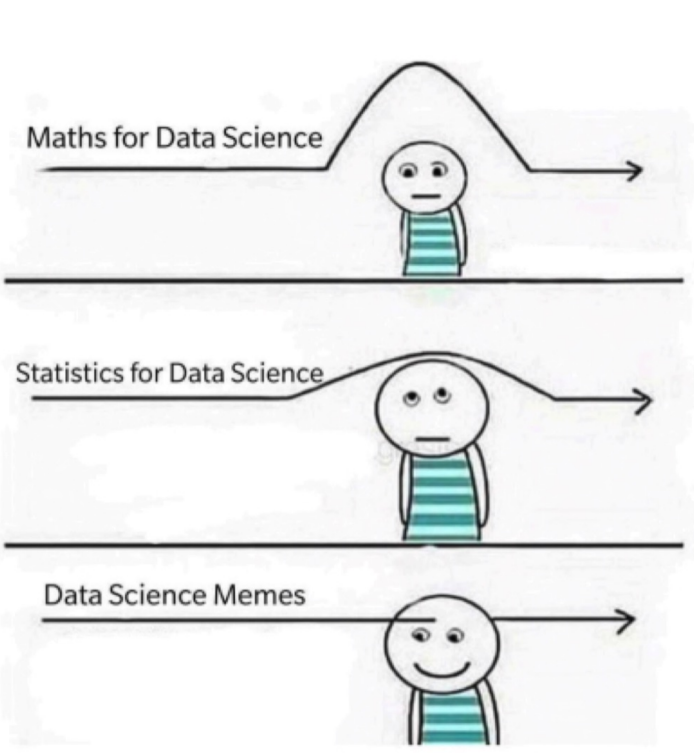
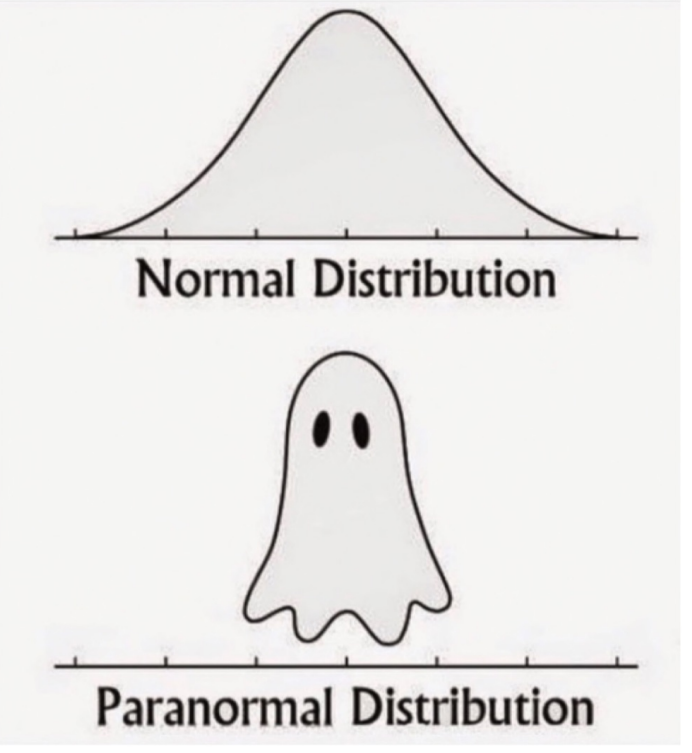
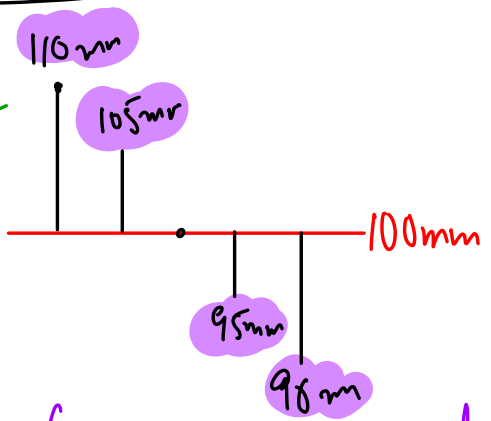
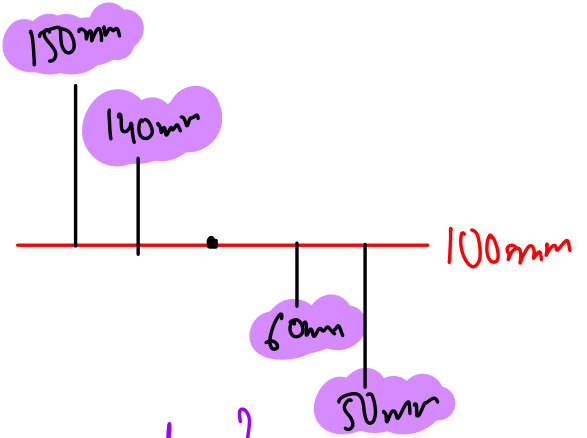
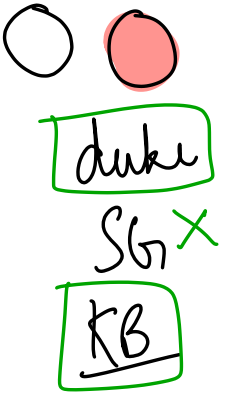

Maths for Data Science
Statistics for Data Science
Data Science Memes


Normal Distribution
Paranormal Distribution

India v/s WI







How far are these values from mean values? on an avg:

$110 - 100 \rightarrow 10$

$105 - 100 \rightarrow 5$

$100 - 100 \rightarrow 0$

$95 - 100 \rightarrow -5$

$90 - 100 \rightarrow -10$

$$\frac{10+5+0-5-10}{5} = 0$$

Absolute

$$\frac{10+5+0+|-5|+|-10|}{5} = 6$$

Square

$$(10^2+5^2+0^2+(-5)^2+(-10)^2)/5$$

$$=(100+25+0+25+100)/5$$

$$= 50$$

$$\frac{\text{Sum of Squares of errors}}{n} = MSE \left[\frac{\sum (O-E)^2}{n}\right]$$

$$\frac{\text{Sum of Absolute errors}}{n} = MAE$$

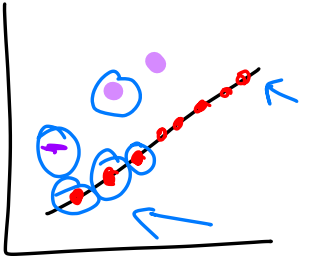
Penalty

Precision
Recall
F1 Score
ROC AUC Curve
RMSE
MAPE

$$\frac{\sum (\text{Observed} - \text{mean})^2}{n} = \text{Variance}$$

$$\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} = \text{Variance}$$

$\bar{x}$, μ = Mean



Variance $\dfrac{(\quad)_{mm} - (\quad)_{mm})^2}{n} \Rightarrow mm^2$

$$\text{Variance} = \sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \quad mm^2$$

$$\Rightarrow \text{Std} = \sqrt{\text{Variance}}$$

$$(\sigma) = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

$\Rightarrow$ Std $\Rightarrow \sigma = 7.07\,mm$

M1

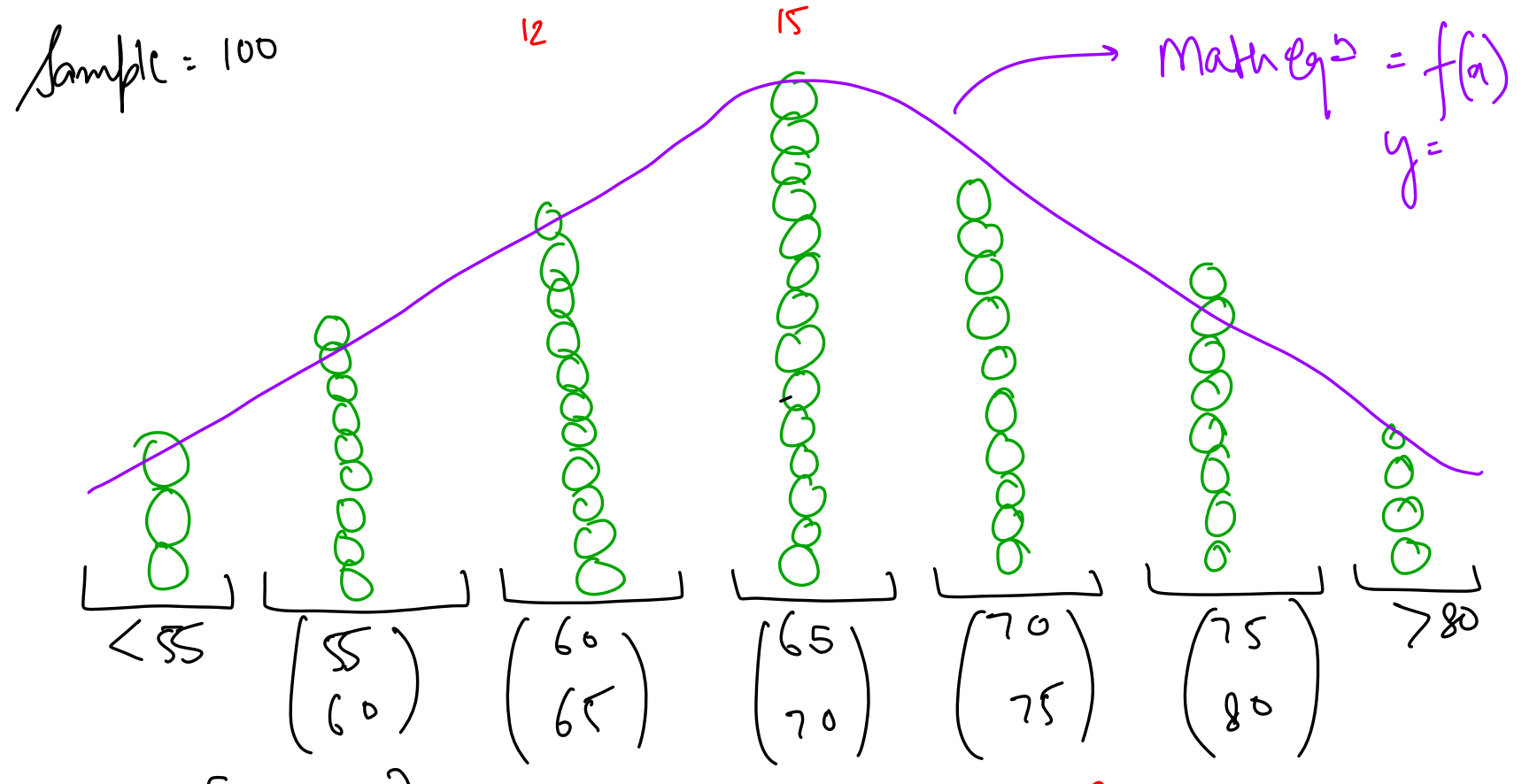
$\sigma^2 = 50\,mm^2$

$\sigma = 7.07\,mm$

M2

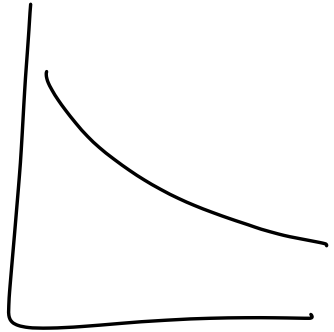
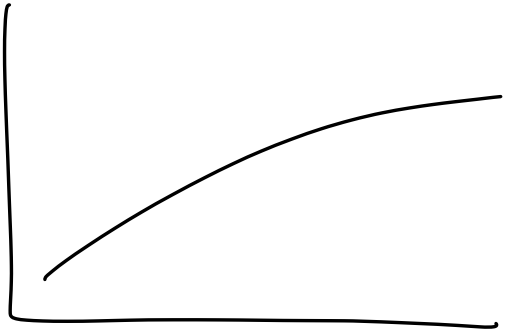
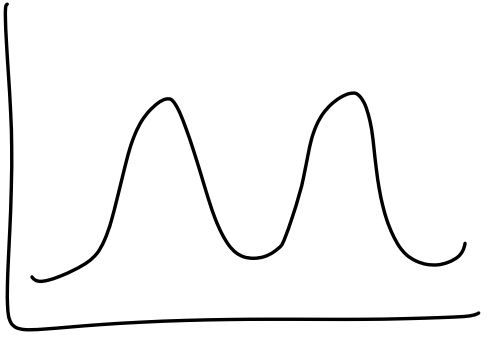
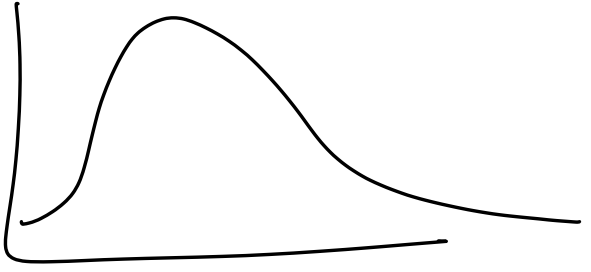
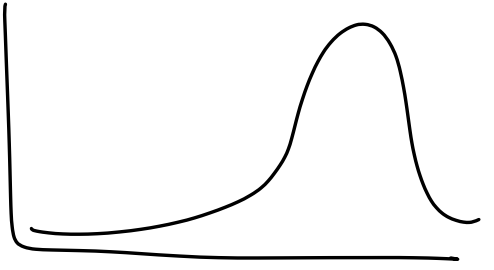
$\sigma^2 = 1640\,mm^2$

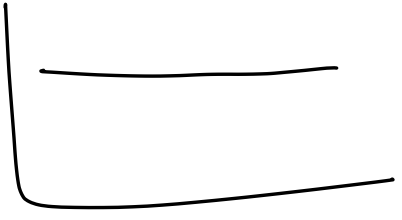
$\sigma = 40.49\,mm$



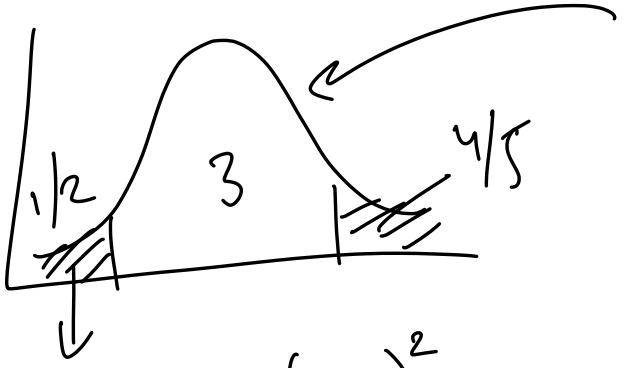
$P[X < 55] = \frac{3}{100}$

$P[60 < X < 70] = \frac{27}{100}$



Appraisal



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

IQ

90-110

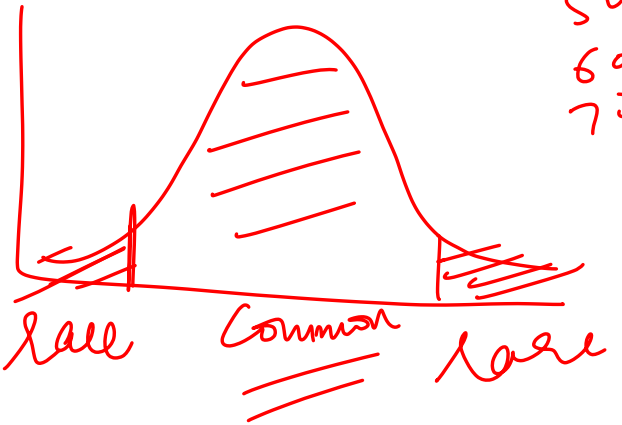
150

170

50

60

70



VARIABLE

DISCRETE

0, 1, 2, 3, 4, 5

CONTINUOUS

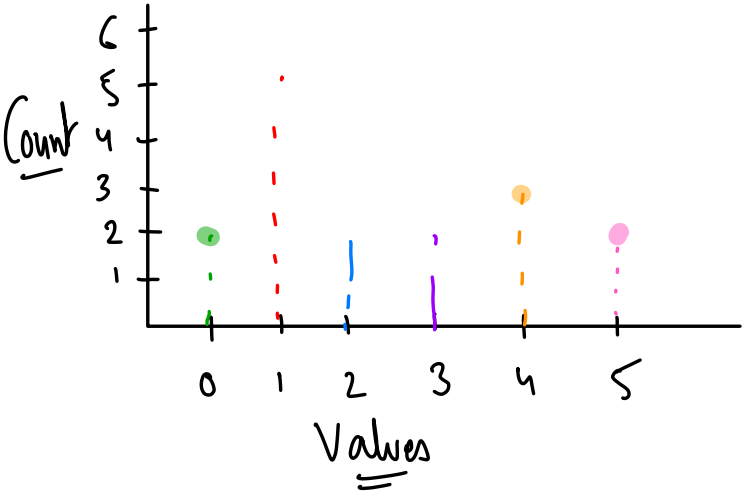
64.5, 64.92, 72.34, 54.67

→ Normal

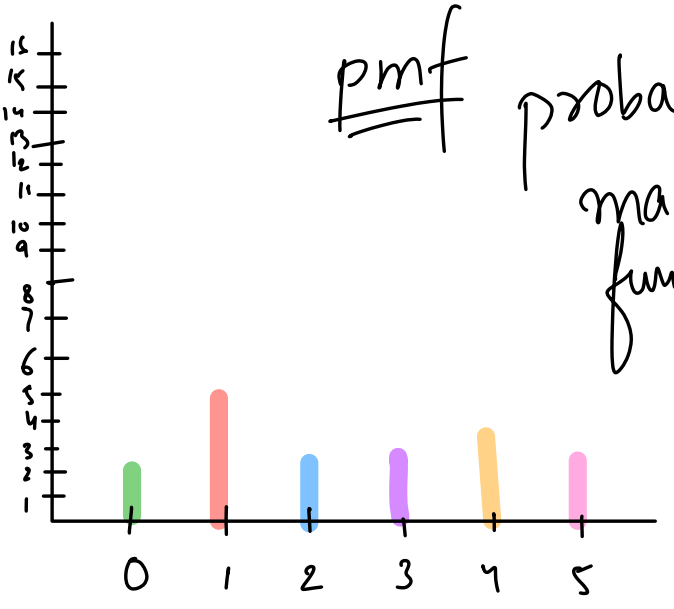
→ Gaussian

→ Bell Curve.

DISCRETE : 0, 1, 2, 3, 4, 5, 0, 1, 1, 1, 1, 2, 3, 4, 5, 4



$\Rightarrow$ density (/16)

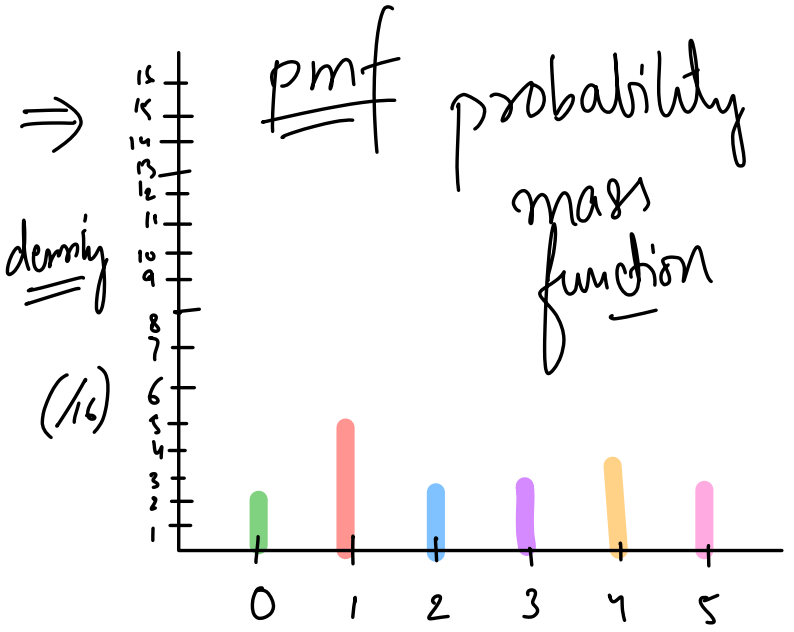


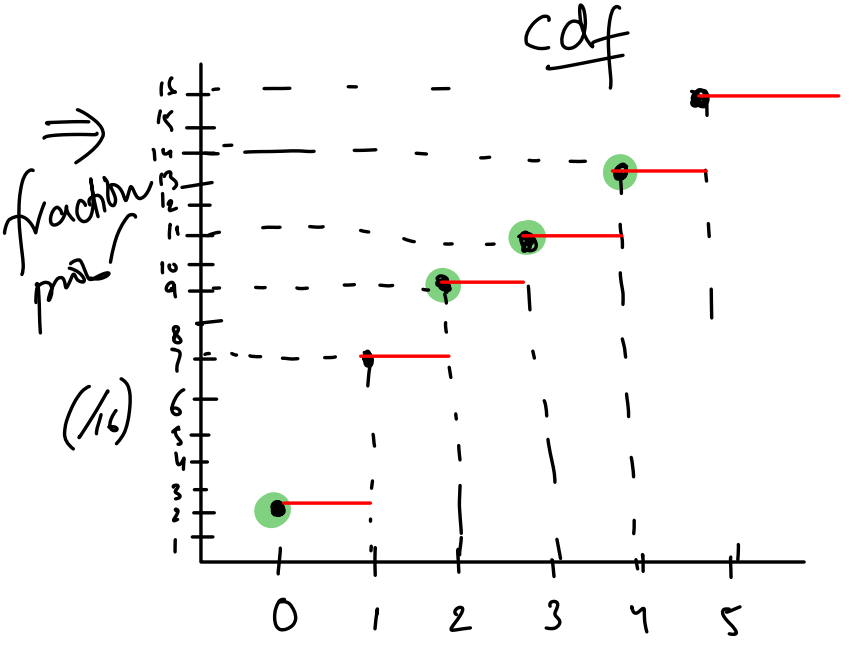
pmf probability mass function

$P[X=0] = 2/16$ $P[X=3] = 2/16$

$P[X=1] = 5/16$ $P[X=4] = 3/16$

$P[X=2] = 2/16$ $P[X=5] = 2/16$





$P[X=0] = 2/16$ $P[X=3] = 2/16$

$P[X=1] = 5/16$ $P[X=4] = 3/16$

$P[X=2] = 2/16$ $P[X=5] = 2/16$

$P[X=0] = 2/16$ 0

$P[X \le 1] = 7/16$ 0+1

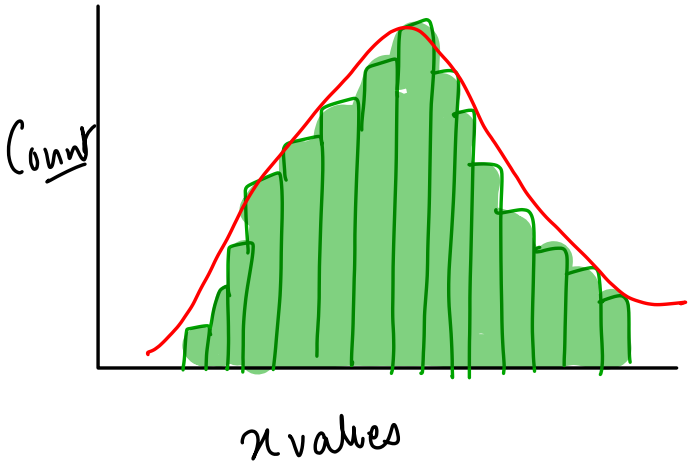
$P[X \le 2] = 9/16$ 0+1+2

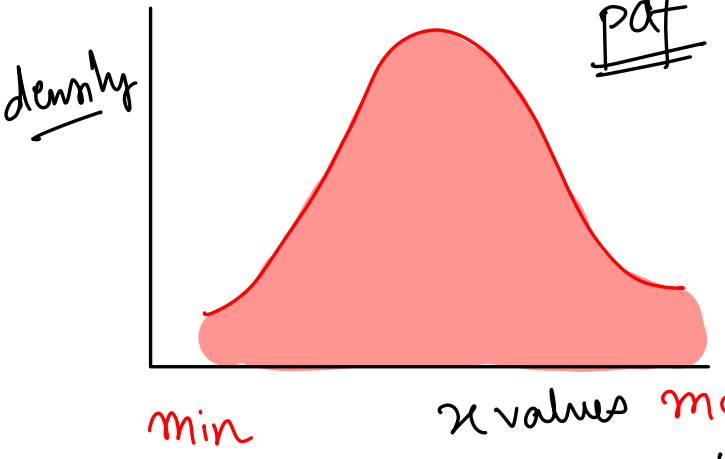
0+1+2+3

0+1+2+3+4

CDf

CONTINCOUS : 79.04, 49.05, 52.045781 . . . .





pdf

Integration

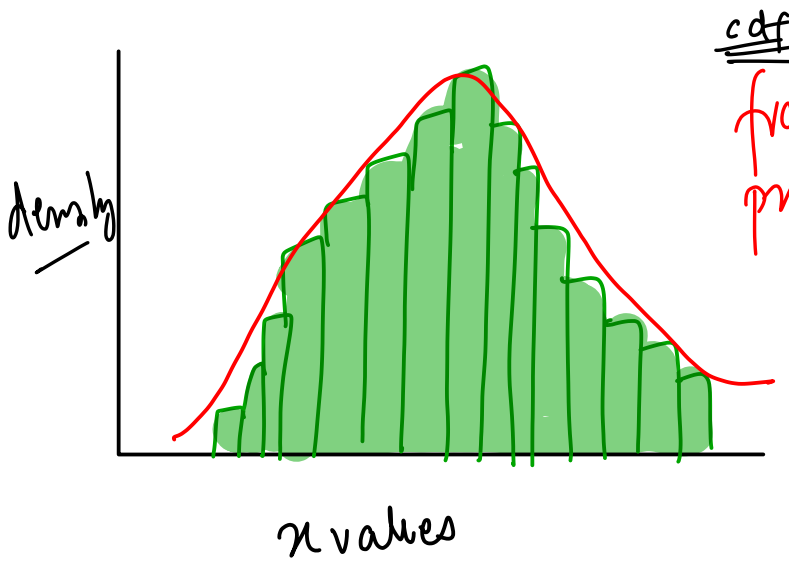
$b(x)$

probability density function

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

$$P[X = 59.2345678]$$

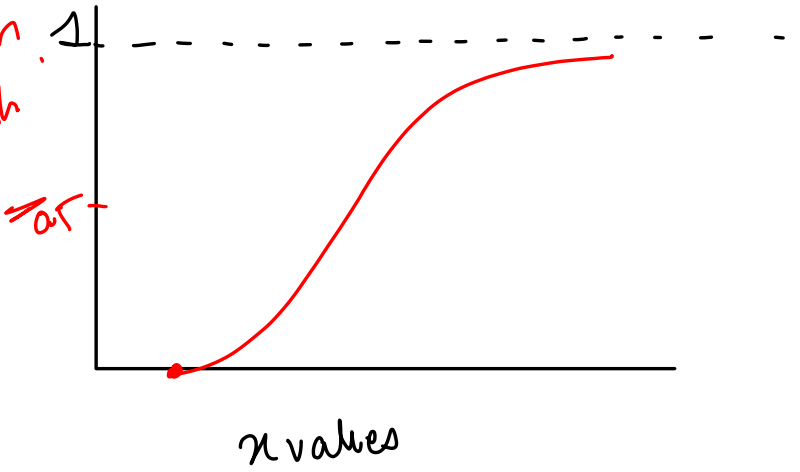
$$P[y_1 < X < y_2] \; P[X < y] \approx 0$$



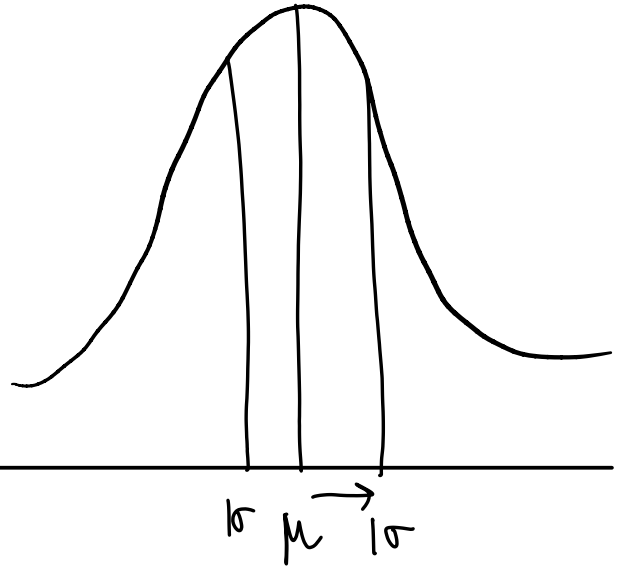
cdf

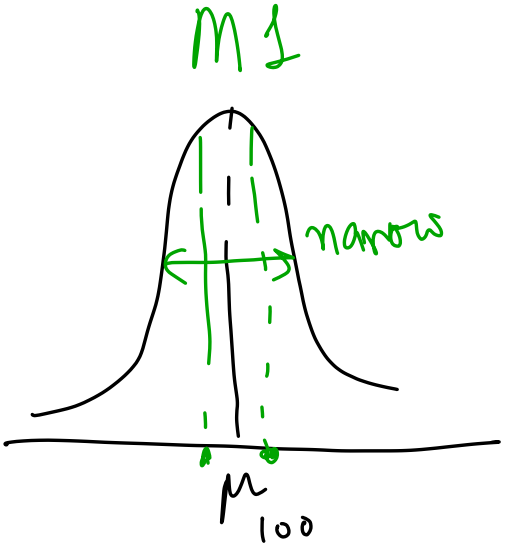
fraction probability

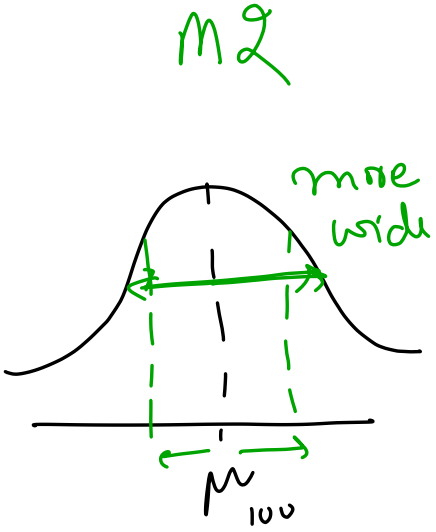
or



prob







$\sigma_{M_1} < \sigma_{M_2}$

$\sigma^2_{M_1} < \sigma^2_{M_2}$

68/95/99 Rule → Gaussian Distribut