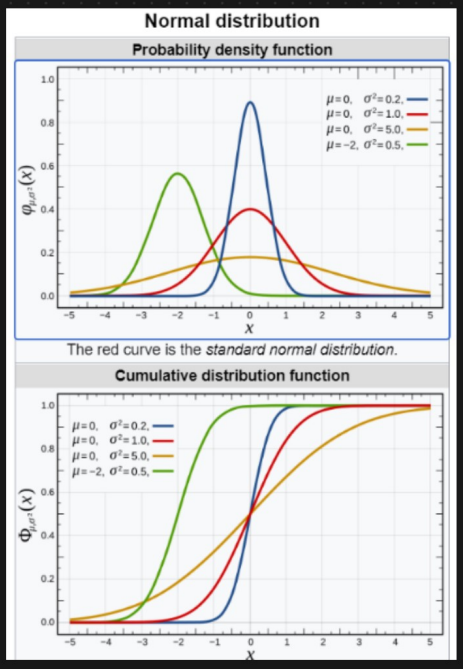
Normal / Gaussian Distribution

In probability theory and statistics, a normal distribution or Gaussian distribution is a type of continuous probability distribution for a real-valued random variable



① Continuous random variable (pdf)

$X = \{ - - - - \}$

$\sigma^2 \uparrow\uparrow$

Spread $\uparrow\uparrow$

⇒ Symmetric distribution

M = median = mode

Eg: Weights of students in a class {Doctors}

Heights of students in a class {Doctors}

IRIS DATASET → Petal length, Sepal Length

Petal Width Sepal Width

Notation $N(\mu, \sigma^2)$

Researchers

Parameters : $\mu \in R$ = mean

$\sigma^2 \in R > 0$ = Variance

$x \in R$.

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

Mean of Normal / Gaussian

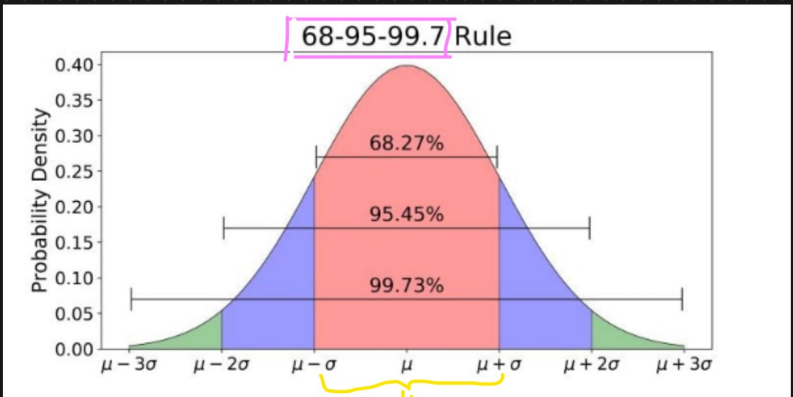
$$\mu = \sum_{i=1}^{n} \frac{x_i}{n}$$

Variance

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

$\sigma = \sqrt{\text{variance}}$

Empirical Rule of Normal / Gaussian Distribution



QQ plot

$X = \{ - \quad - \quad - \quad - \quad - \quad \cdots \}$ => Normal / Gaussian Distribution

- Probability

$$Pr(\mu-\sigma \leq X \leq \mu+\sigma) \approx 68\%$$

$$Pr(\mu-2\sigma \leq X \leq \mu+2\sigma) \approx 95\%$$

$$Pr(\mu-3\sigma \leq X \leq \mu+3\sigma) \approx 99.7\%$$