# PROBABILITY

① $P = \dfrac{\text{Required}}{\text{Total}}$

② $0 \le P(A) \le 1$
- $0$ ↳ Null/Impossible event
- $1$ ↳ sure event

③ $P(\bar{A}) = 1 - P(A)$ $\quad(\bar{A}/A'/A^c)$

④ $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

⑤ $P(A\cup B\cup C) = P(A) + P(B) + P(C) - P(A\cap B) - P(B\cap C) - P(A\cap C) + P(A\cap B\cap C)$

⑥ $P(\overline{A\cup B}) = P(\bar{A}\cap\bar{B}) = 1 - P(A\cup B)$

⑦ $P(\overline{A\cap B}) = P(\bar{A}\cup\bar{B}) = 1 - P(A\cap B)$

⑧ $P(\text{only } A) = P(A\cap\bar{B}) = P(A) - P(A\cap B)$

* **COIN TOSS** (Independent event)
'n' coin is tossed $\Rightarrow n(S) = 2^n$

⑨ $P(\text{both } A \,\&\, B) = P(A\cap B)$

⑩ $P(\text{at least one}) = P(\text{either } A \text{ or } B) = P(A\cup B)$
$= 1 - P(\text{no one})$

⑪ $P(\text{neither } A \text{ nor } B) = P(\bar{A}\cap\bar{B}) = P(\overline{A\cup B})$

⑫ $P(\text{exactly one}) = P(A\cap\bar{B}) + P(\bar{A}\cap B)$

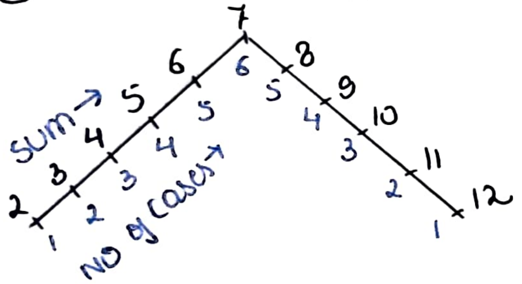
⑬ Mutually Exclusive events: can't occur together
$P(A\cap B) = 0$ ; $A\cap B = \phi$

⑭ Independent Events: occurance of 1 doesn't affect other
$P(A\cap B) = P(A)\cdot P(B)$

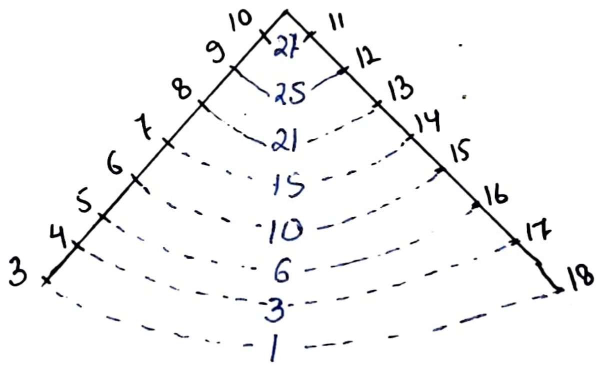
⑮ Dependent Events:
$P(A\cap B) = P(A)\cdot P(B/A)$
$= P(B)\cdot P(A/B)$

* **DICE THROW** (Independent event)
'n' dice throw $\Rightarrow n(S) = 6^n$

⑯ 2 Dice rolled




⑰ 3 Dice rolled




⑱ If no. of coin > 3 , no. of dice > 2 ⇒ use Binomial dist.

**Pack of card (52)**
- Red (26)
  - Heart (13) ♡
  - Diamond (13) ◇
- Black (26)
  - Spade (13) ♤
  - Club (13) ♧

⑲ Each suit contains
(2, 3, 4, 5, 6, 7, 8, 9, 10) → Number cards = 9
J, K, Q → Face cards = 3
J, K, Q, A → Honour cards = 4

⑳ If A & B are independent events
- $\bar{A}$ & $B$ " " "
- $A$ & $\bar{B}$ " " "
- $\bar{A}$ & $\bar{B}$ " " "

㉑ By default → without Replacement.

㉒ $P(A/B) = \dfrac{P(A\cap B)}{P(B)}$

㉓ $P(A/B) + P(\bar{A}/B) = 1$
$P(A/\bar{B}) + P(\bar{A}/\bar{B}) = 1$

* $P(A) = P(A/x)P(x) + P(A/y)P(y)$
* If $x$ is average, there must be atleast 1 data point $\ge x$ & $\le x$.

(24) A, B, C are pair wise independent events
$P(A \cap B) = P(A) \cdot P(B)$
$P(B \cap C) = P(B) \cdot P(C)$
$P(A \cap C) = P(A) \cdot P(C)$
↳ All possible pairs

(25) A, B, C are mutually independent events
$P(A \cap B) = P(A) \cdot P(B)$
$P(B \cap C) = P(B) \cdot P(C)$
$P(C \cap A) = P(C) \cdot P(A)$
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$
↳ All possible combinations of A, B & C

(26) Total no. of conditions for mutually independence of n events $= 2^n - 1 - n$

(27) Mutually independent events are pair wise independent but not vice-versa.

PDF: Prob. density/Mass fuⁿ
CDF: cumulative distribution fuⁿ

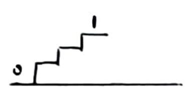
## RANDOM VARIABLE

↗ value of x is discrete set of nos.
* Discrete Random variable ($\Sigma$)

P(x) is PDF.

PDF: $\sum_{-\infty}^{\infty} P(x) = 1$

CDF: $F(X) = \sum_{-\infty}^{x} P(x)$



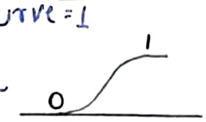
↗ value of x is given in range
* Continuous Random variable ($\int$)

f(x) is PDF

PDF: $\int_{-\infty}^{\infty} f(x)\,dx = 1$

Area under curve = 1

CDF: $F(X) = \int_{-\infty}^{x} f(x)\,dx$



(1) CDF: $F(x) = P(X \le x)$ ↗F(x)
$F(-\infty) = 0$
$F(\infty) = 1$

(2) $P(X=2) = F(X=2) - F(X=1)$
(in discrete RV)

(3) CDF is continuous fuⁿ
(for cont. RV)
CDF is discontinuous fuⁿ
(for discrete RV)

(4) $0 \le F(x) \le 1$

(5) CDF ↔ PDF ; both r interconvertible.

(4) PDF: $f(x) = \frac{d}{dx} F(X)$ ↗ f(x)/P(x)

(5) $P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)\,dx$
(for cont RV)

(6) In cont. RV
$P(a < x < b) = P(a \le x < b)$
$= P(a < x \le b) = P(a \le x \le b)$

(7) In discrete RV all are different
$P(0 < x < 2) = P(1)$
$P(0 < x \le 2) = P(1) + P(2)$

### 2D Random variables

- Discrete RV
$\sum_{y=} \sum_{x=} P(x,y) = 1$
- continuous RV
$\int_y \int_x f(x,y)\,dx\,dy = 1$

### Marginal PDF

- eliminate other
Given $f(x,y)$.
- $f(x) = \sum_{y=} f(x,y)$ (in discrete)
Range of x must be constant
- $f(x) = \int_{y=} f(x,y)\,dy$ (cont.)
- X & Y are independent if $f(x,y) = f(x) \cdot f(y)$

### Mean/Expectation/Average

(1) $E(X) = \sum x P(x)$ (Discrete RV)

(2) $E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$ (cont. RV)

(3) $-\infty < E(x) < \infty$

(4) $E(g(x)) = \sum g(x) P(x)$ (discrete)
$= \int_{-\infty}^{\infty} g(x) f(x)\,dx$ (cont.)

(5) $E(C) = C$

(6) $E(aX) = aE(X)$

* $E(X) = E_1 P(A) + E_2 (P(A^c))$
For ques like coins r tossed until HH occurs use recursion

(7) $E(aX+b) = aE(X) + b$

(8) $E(X+Y) = E(X) + E(Y)$

(9) $E(X-Y) = E(X) - E(Y)$

(10) $E(aX^n) = aE(X^n)$

(11) $E(XY) = E(X) \cdot E(Y/X)$
If X & Y are independent
↳ $E(XY) = E(X) \cdot E(Y)$ ↲ may or not

(12) $E(X^2) \ne (E(X))^2$

(13) $E(1/X) \ne 1/E(X)$

### Standard Deviation ($\sigma$)

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

$\sigma$ is always +ve i.e. $\ge 0$

## Variance

1. $V(X) = E(X^2) - [E(X)]^2 = E[(x - E(X))^2]$
2. Its always +ve. i.e. $\geq 0$.
3. $V(C) = 0$
4. $V(aX) = a^2 V(X)$
5. $V(-X) = V(X)$
6. $V(aX+b) = a^2 V(X)$
7. $V(X+Y) = V(X) + V(Y) + 2\,\text{cor}(x,y)$ — correlation b/w $x$ & $y$

   $V(X-Y) = V(X) + V(Y) - 2\,\text{cor}(x,y)$
8. If $X$ & $y$ are independent RV, $\text{cor}(x,y) = 0$

   $V(X+Y) = V(X) + V(Y)$

   $V(X-Y) = V(X) + V(Y)$
9. By default consider independent

$E(e^{tx}) = 1 + tx + \frac{t^2x^2}{2!} + \frac{t^3x^3}{3!} + \ldots$

## Moment Generating fu^n:

$E(X) \rightarrow$ 1st moment = Mean

$E(X^2) \rightarrow$ 2nd moment

$$E(X^n) = \left[\frac{d^n}{dt^n} E(e^{tx})\right]_{t=0}$$

## PROBABILITY DISTRIBUTION

| Discrete dist. | Cont. dist. |
|---|---|
| • Binomial | • Uniform |
| • Bernoulli | • Exponential |
| • Hypergeometric | • Normal |
| • Poisson | |
| • Geometric | |

* Any problm which can be solved by Binomial can also be solved by Poisson.

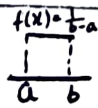
| Binomial Distribution | Bernoulli Distr. | Hypergeometric Distr. | Poisson Distr. ($e = 2.718$) | Geometric Dist. |
|---|---|---|---|---|
| • Discrete RV<br>• $P+q=1$<br>• Independent events | } same<br>• $n=1$ | • Discrete RV<br>• $p+q=1$<br>• $P$ & $q$ not independent | • Discrete RV<br>• $P+q=1$<br>• Independent events<br>• $n \to \infty,\ p \to 0$ | } same<br>• No combination position is fixed |
| $P(X) = {}^nC_x\, p^x q^{n-x}$ | $P(X) = p^x q^{1-x}$ | • $P = \frac{\text{Required}}{\text{Total}}$<br>• without replacement<br>• If we take with replacement, $p$ & $q$ will become independent hence its binomial dis. | • $P(X) = \frac{e^{-\lambda}\lambda^x}{x!}$<br>$X \geq 0,\ \lambda = np$ | • $P(X) = p^1 q^{X-1}$ |
| $E(X) = np$ | $E(X) = p$ | | • $E(X) = V(X) = \lambda$ | • $E(X) = \frac{1}{p}$ |
| $V(X) = npq$ | $V(X) = pq$ | | ($X$ & $\lambda$ should talk about same thing) success or failure | • $V(X) = \frac{q}{p^2}$ |

$n \to$ No. of trials, $X \to$ No. of success required

$P \to$ Prob of success in 1 trial, $q$ = Prob of failure in 1 trial.

### Uniform dist

PDF:



$$f(X) = \frac{1}{b-a}$$

$a \leq x \leq b$

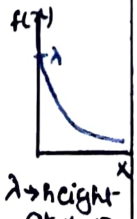
$E(X) = \frac{a+b}{2} = \int x f(x)\,dx$

$V(X) = \frac{(a-b)^2}{12}$

### Exponential dist

$f(x) = \lambda e^{-\lambda x}$

$x \geq 0$



$\lambda \to$ height at $x=0$

$E(X) = \frac{1}{\lambda}$

$V(X) = \frac{1}{\lambda^2}$

$E(X^n) = \frac{n!}{\lambda^n}$

### Normal/Gaussian dist

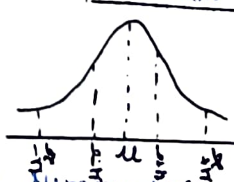
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$f(x)$ is max at $x = \mu$

$f_{max} = \frac{1}{\sqrt{2\pi\sigma^2}}$

Mean $= \mu$

SD $= \sigma$



• Bell shaped curve symmetric about $x$ = mean

$$\int_{-\infty}^{\mu} f(x)\,dx = \int_{\mu}^{\infty} f(x)\,dx = \frac{1}{2}$$

## Z distr (Standard Normal dist.)

TO solve prob ques, normal curve is converted to Z distr

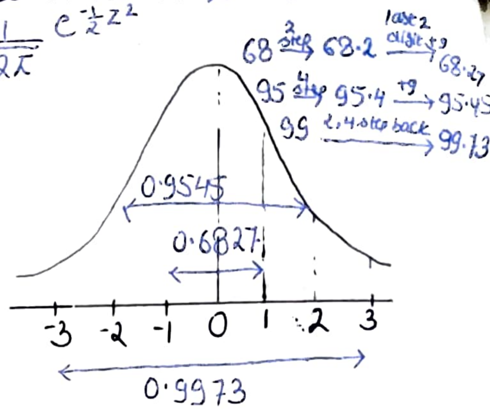
① $z = \frac{x-\mu}{\sigma}$

② $P(Z>0) = P(Z<0) = \frac{1}{2}$

③ $\mu = 0$ , $\sigma = 1$ (on comparing)

④ If X & Y are independent Normal random variables then $(aX \pm bY)$ is also a normal random variable

⑤ $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$




---

## STATISTICS

* In terms of prob, median is a point m such that $P(-\infty < X < m) = P(m < X < \infty)$

### A. MEAN

① Ungrouped data

$\text{Mean} = \frac{\sum x_i}{n}$

② Grouped data

$\text{Mean} = \frac{\sum x_i f_i}{\sum f_i}$

③ If $x_i$ given in interval find mean of class interval

### B. MEDIAN

① ungrouped data

- Arrange in Asc/Des order
- If n is odd, median $= \frac{(n+1)}{2}$ th term
- If n is even, median = mean of $\frac{n}{2}$th, $(\frac{n}{2}+1)$th term

② Grouped data use cumulative freq.

③ If $x_i$ given in interval & N is even

$\text{Median} = l_m + \frac{(\frac{N}{2} - PCF)}{f_m} h$

$l_m$ → lower limit of median class (N/2 cum. freq lie in which class)

PCF → Prev cum freq

$f_m$ → freq of median class

$h$ → height of class

### C. MODE

$x_i$ having highest freq.

NOTE: {1, 2, 3, 4} NO Mode

① If $x_i$ is given in range

$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) * h$

$l$ → lower limit of modal class

$f_1$ → freq of modal class

$h$ → height of modal class

- Dean = Dian (not always)

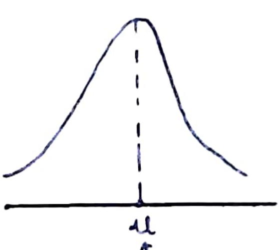
1 Mode + 2 Mean = 3 Median

- SD, $\sigma = \sqrt{E(x^2) - (E(x))^2}$

$\sigma = \sqrt{\frac{\sum x^2 f}{\sum f} - \left(\frac{\sum x f}{\sum f}\right)^2}$

---

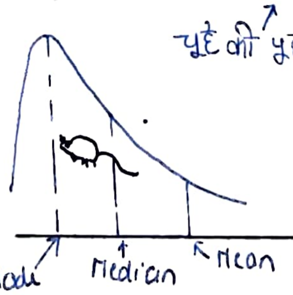
**Normal dist.**




Mean = Median = Mode

Right skewed / +vely skewed

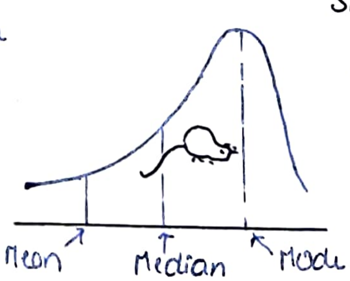
पूँछ की पूँछ +ve side




Mean ≥ Median ≥ Mode

Left skewed / -vely skewed

पूँछ -ve side




Mode ≥ Median ≥ Mean

---

- coeff of variation $= \frac{\sigma}{\mu}$ ← SD / ← Mean
- It gives amount of fluctuation in data.

NOTE: In Normal dist, if we have to find prob at a point

$P(X=16) = P(15.5 < X < 16.5)$ ← Find this area

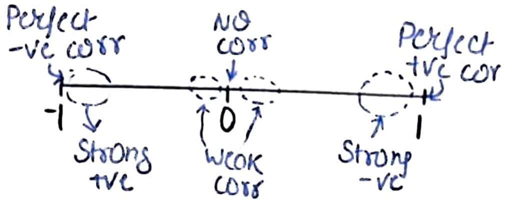
## Correlation:

$r$ : correlation coefficient / index → tells nature & strength of rel$^n$ b/w data set

① $r \in [-1, 1]$

② Direct / +ve corr. $r = 0$ to $1$ $\qquad x\uparrow \mid x\downarrow$ ; $y\uparrow \mid y\downarrow$

③ Inverse / -ve corr. $r = -1$ to $0$ $\qquad x\uparrow \mid x\downarrow$ ; $y\downarrow \mid y\uparrow$




④ $r = \dfrac{n\Sigma xy - \Sigma x \Sigma y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$

⑤ $r = \dfrac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y}$

$\mathrm{cov}(X,Y) = \dfrac{\Sigma(X-\bar{X})(Y-\bar{Y})}{n}$
$= E[(X-E(X))(Y-E(Y))]$

$X$ & $Y$ are two data sets

**NOTE**: AGP fibonacci series
$1\cdot x^0 + 1\cdot x^1 + 2\cdot x^2 + 3\cdot x^3 + 5\cdot x^4 + 8\cdot x^5 + \cdots = \dfrac{1}{1-x-x^2}$

## Regression:

① $X = f(Y)$ — $X$ on $Y$ $\qquad Y = f(X)$ — $Y$ on $X$

M1 ② $Y$ on $X$

$Y = a + bX$

do $\Sigma$ → $\Sigma Y = na + b\Sigma X$ —①

$\times X$ do $\Sigma$ → $\Sigma XY = a\Sigma X + b\Sigma X^2$ —②

From eq ① & ② find $a$ & $b$

③ Dividing eq① by $n$
$\bar{y} = a + b\bar{x}$
$\bar{x}$ & $\bar{y}$ (mean) also lie on regression line.

$n \rightarrow$ No. of ele in a data set

M2 ④ $Y$ on $X$
y pahle x baad me

$(y-\bar{y}) = b_{yx}(x-\bar{x})$

$b_{yx} = \dfrac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$

last me → x so deno me $\Sigma x$

⑤ $b_{yx}$ is same as $b$ in Method 1

⑥ $X$ on $Y$
x phle y baad me

$(x-\bar{x}) = b_{xy}(y-\bar{y})$

$b_{xy} = \dfrac{n\Sigma xy - \Sigma x\Sigma y}{n\Sigma y^2 - (\Sigma y)^2}$

last me y so denom me $\Sigma y$

⑦ Rel$^n$ b/w correlation & Regression coeff.

(i) $b_{xy} = r\dfrac{\sigma_x}{\sigma_y}$

$b_{yx} = r\dfrac{\sigma_y}{\sigma_x}$

sign same

(ii) $b_{xy} * b_{yx} = r^2$

⇒ **convolution theorem**

Let $X, Y$ be random var & $P_X, P_Y$ be PDF.
$Z = X + Y$

- $P_Z(z) = \sum_x P_X(x) P_Y(z-x)$ (for discrete RV)
- $f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx$ (for continuous RV)

* If $X_1$ and $X_2$ are independent exponential random var. with parameters $\lambda_1$ & $\lambda_2$ then $X = \min(X_1, X_2)$ is expo. RV with parameter $\lambda_1 + \lambda_2$.

* A ball is picked from bag & placed back along with 1 more ball of same colour every trial, Then, P(red ball) is same in every iteration.
(GATE 21)