PROBABILITY

→ Classical approach - $P(E) = \frac{h}{n}$ → fav o/c

→ Rel. Freq approach - $P(E) = \lim_{n \to \infty} \frac{h}{n}$

→ Axiomatic approach - ① $P(A) \geq 0$ ② $P(S) = 1$

③ $P(\bigcup_i A_i) = P(A_1) + P(A_2) + \ldots$

→ $P(\phi) = 0$

$P(\bar{A}) = 1 - P(A)$

$A \subseteq B$, then $P(A) \leq P(B)$.

$P(B-A) = P(B) - P(A)$

$0 \leq P(A) \leq 1$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Add^n Rule)

→ Conditional probability - $P(A|B) = \frac{P(A \cap B)}{P(B)}$

→ Multiplication rule - $P(A_1 \cdot A_2 \ldots A_n) =$

$P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 A_2) \ldots P(A_n|A_1 A_2 \ldots A_{n-1})$

→ Baye's Rule - $P(A) = P(A|B) \cdot P(B) + P(A|\bar{B}) P(\bar{B})$

→ Baye's Rule - $P(B_k|A) = \frac{P(B_k) \cdot P(A|B_k)}{P(B_1) \cdot P(A|B_1) + \ldots + P(B_n) \cdot P(A|B_n)}$

→ Independent events - $P(A \cap B) = P(A) \cdot P(B)$.

→ Random Variable - A r.v is a func^n that assigns to each outcome $\omega \in S$ a real no.

A real valued func^n $X: \Omega \to R$ → r.v

Ω → event space.

→ PMF - $F(x) = P[X \leq x] = \sum_{x_i < x} f(x_i)$ $\boxed{\sum_{x_i} f(x_i) = 1}$

PDF - $\int_a^b f(x) \cdot dx = P[a < X < b]$

$F(a) = \int_{-\infty}^{a} f(x) \cdot dx$ $P[X \leq a]$

$$\boxed{\int_{-\infty}^{\infty} f(x) \cdot dx = 1}$$

→ Relat^n btw Distribution Func^n & Density Func^n

$\frac{d}{dx} F(x) = f(x)$ $F(x) = \int_{-\infty}^{x} f(x)\, dx = P(X \leq x)$

→ Mean - (eg- centre of gravity) $E(X) = \begin{cases} \sum_x x \cdot P(x) & \text{discrete} \\ \int_{-\infty}^{\infty} x \cdot f(x)\,dx & \text{continuous} \end{cases}$

① $E(KX) = K \cdot E(X)$ ② $E(X \pm Y) = E(X) \pm E(Y)$

→ Variance (eg- moment of inertia) $Var(X) = E(X-\mu)^2 = E(X^2) - E(X)^2$

① $Var(const) = 0$ ② $Var(KX) = K^2 \cdot Var(X)$ ③

③ $Var(X+Y) = Var(X) + Var(Y) + 2[E(X-\mu_x)(Y-\mu_y)]$

→ For a continuous r.v X with cdf F & pdf f

$$E(X) = \int_0^{\infty} [1-F(t)]\,dt - \int_0^{\infty} F(-t)\cdot dt = \int_0^{\infty} P(X>t)\,dt - \int_0^{\infty} P(X\leq -t)\,dt$$

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

$$\left.\begin{aligned} E(X) &= \int_0^{\infty} (1-F(t))\cdot dt \\ E(X^2) &= 2\int_0^{\infty} t\,[1-F(t)]\cdot dt \end{aligned}\right\}$$

X is not continuous
X is a non-neg r.v.

→ SPL ~~PRO~~BABILITY DISTRIBUTIONS

① Discrete Uniform R.V

$$f(x) = \begin{cases} \frac{1}{b-a+1} & a \leq x \leq b \\ 0 & o/w \end{cases} \qquad F(x) = \begin{cases} 0 & x < a \\ \frac{x-a+1}{b-a+1} & a \leq x \leq b \\ 1 & x > b. \end{cases}$$

$E(X) = \frac{a+b}{2}$ $Var(X) = \frac{n^2-1}{12}$; $n = b-a+1$

② Bernoulli RV

Pmf -

X	0	1
P(x)	q	p

Mean - $E(x) = p$
$Var(X) = pq$

③ BINOMIAL R.V

Pmf $P(x) = {}^nC_x\, p^x \cdot q^{n-x}$ $x = 0, 1, 2 - - n$

$E(X) = np$ $Var = npq$

④ Geometric Distribution

$P(X) = q^{x-1} \cdot p$

Mean $= 1/p$ $Var = q/p^2$

Memoryless property of GD

$$P[X > m+n \mid X > m] = q^n$$

⑤ Poisson Distribution -

PMF - $f(x) = \dfrac{e^{-\lambda} \cdot \lambda^x}{x!}$ mean = var = λ

→ SPL CONTINUOUS DIST

① Continuous Uniform Distribution

Pdf $f(x) = \begin{cases} \dfrac{1}{b-a} & x \in (a,b) \\ 0 & o/w \end{cases}$

cdf $F(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a < x < b \\ 1 & x \geq b \end{cases}$

$$P[k_1 < X < k_2] = \frac{k_2 - k_1}{b-a}$$

$E(X) = \dfrac{a+b}{2}$ $Var(X) = \dfrac{(a-b)^2}{12}$

② Exponential Distribution (Time per single event)

Mean = $1/\lambda$ Var = $1/\lambda^2$

Memoryless property of exponential dist -

$$P[X > t+s \mid X > s] = P[X > t]$$

③ Weibull Distribution -

$f(x) = \alpha \beta x^{\beta - 1} \cdot e^{-\alpha x^{\beta}}$ α - scale parameter, β - shape parameter

Mean = $\dfrac{1}{\alpha^{1/\beta}} \Gamma\left(\dfrac{1}{\beta} + 1\right)$

$E(X^2) = \dfrac{1}{\alpha^{2/\beta}} \Gamma\left(2/\beta + 1\right)$

④ Normal | Gaussian Distribution -

PDF = $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-1/2 \left(\frac{x-\mu}{\sigma}\right)^2}$ $-\infty < x < \infty$

Moments

→ Let X be the r.v with mean μ $(E(x)-\mu)$

$E(x^r)$ - r^{th} moment of X (r^{th} moment abt origin)

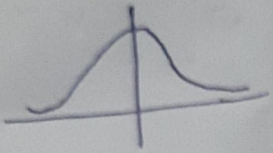
$E(X-\mu)^r$. r^{th} central moment of x (r^{th} moment abt mean)

$E(x-\mu)^r - \mu_r$

$r=1$ $E(X)$ → mean

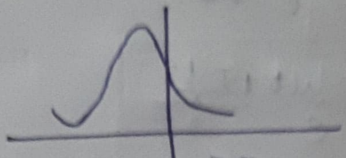
$r=2$ $E(X-\mu)^2$ → variance.

→ $\frac{\mu_3}{\sigma_3}$ measure of symmetry (skewness)



0

symmetric



>0

skewed to right

<0

skewed to left

→ $\frac{\mu_4}{\sigma_4}$ measure of kurtosis (peakness)

: SND ; >3 peak than SND ; <3 less than SND

→ Moment generating function : MGF

$$M_X(t) = E(e^{tX})$$

$\frac{d^r}{dt^r}\left[M_X(t)\right]$ at $t=0$ gives $E(x^r)$

→ Properties of MGF -

1. 2 r.v X & Y with same MGF have the same prob dist (i.e) mgf of a rv is unique.
2. $M_X(0) = 1$
3. $M_{cX}(t) = M_X(ct)$
4. $M_{c+x}(t) = e^{ct} \cdot M_X(t)$
5. $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$

→ Erlang Distribution

Pdf ; $f(x) = \frac{\lambda e^{-\lambda x} \cdot (\lambda x)^{k-1}}{\Gamma(k)}$ $x \geq 0$

$\Gamma(k) = (k-1)!$

parameters λ & k. MGF = $\left(\frac{\lambda}{\lambda - t}\right)^r$

Mean = $\frac{k}{\lambda}$ Var = k/λ^2

Gamma Distribution (Generalisatⁿ of erlang)

Mean = $\frac{r}{\lambda}$ Var = $\frac{r}{\lambda^2}$

Distribution	MGF	$M_x(t) = E(e^{tx})$ x - r.v
① Binomial	$(pe^t + q)^n$	
② Exponential	~~$e^{\lambda(e^t-1)}$~~ $\lambda/\lambda - t$	
③ SND	$e^{t^2/2}$	
④ ND	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$	
⑤ erlang	$\left(\frac{\lambda}{\lambda - t}\right)^r$	

→ JOINT PROBABILITY (2 or more r.v's on same sample space)

$f_{xy}(x,y)$ - Joint pdf $P_{xy}(x,y) = P[X = x, Y = y]$ Joint pmf

$F_{xy}(x,y) = P[X \leq x, Y \leq y]$ - Joint cdf

→ Marginal Dist - individual prob distⁿ of r.v's X & Y.

Marginal Dist of x $P_x(x) = \sum_{y \in Y} P_{xy}(x,y)$

" " " y $P_y(y) = \sum_{x \in X} P_{xy}(x,y)$

PROPERTIES

Pmf	Pdf
① $P_{xy}(x,y) \geq 0$	① $f_{xy}(x,y) \geq 0$
② $\sum_{x \in X} \sum_{y \in Y} (x,y) = 1$	② $\int\int f_{xy}(x,y)\, dy\, dx = 1$

→ If $f_{xy}(x,y)$ is joint pdf of x & Y.

Marginal dist of x - $f_x(x) = \int f(x,y)\, dy$

" " " y - $f_y(y) = \int f(x,y)\, dx$

→ Conditional Distribution

$$f_{Y|X=x}(y|x) = \frac{f(x,y)}{f_X(x)}, \quad f_X(x) \neq 0$$

$$f_{X|Y=y}(x|y) = \frac{f(x,y)}{f(y)}, \quad f_y(y) \neq 0$$

→ $E(Y|X=x) = \begin{cases} \sum_{y \in Y} y \cdot f_{Y|X}(y|x) & x \,\&\, y \text{ are discrete} \\ \int y \cdot f_{Y|X}(y|x)\, dy & x \,\&\, y \text{ are cont.} \end{cases}$

$E(Y^2|X=x) = \begin{cases} \sum_{y \in Y} y^2 \cdot f_{Y|X}(y|x) \\ \int y^2 \cdot f_{Y|X}(y|x)\, dy \end{cases}$

→ $E(E(X|Y)) = E(X)$

→ $E(XY) = \iint xy\, f(x,y)\, dy\, dx$

→ $E[X^j Y^k]$ = jk^{th} joint moment of X & Y.

For $j=k=1$, $E(XY)$ is correlation of X & Y.

$E[(X-\bar{X})^j (Y-\bar{Y})^k] \rightarrow jk^{th}$ joint central moment.

$j=k=1 \Rightarrow E[(X-\bar{X})(Y-\bar{Y})] \rightarrow$ co-variance of X & Y.

→ Covariance – measures the deviations of the points X & Y from $E(X) = \mu_x$ & $E(Y) = \mu_y$.

$X - \mu_x$ & $Y - \mu_y$ → same sign ⟶ co variance (+ve)

" " " → diff sign ⟶ co variance (−ve)

→ $\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$ $\quad \sigma_X = SD(X)$, $\sigma_Y = SD(Y)$

co-eff of co-relat^n b/w X & Y

→ $Var(aX+bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\, cov(X,Y)$

$$r_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

→ Mean and Variance of sum of r.v's

$$Var(W_n) = \sum_i Var(X_i) + 2 \sum_{i<j} covar(X_i, X_j)$$

→ If $X_1, X_2 \ldots X_n$ are pairwise independent or pairwise corelated then

$$Var(W_n) = \sum_i Var(X_i)$$

→ LIMIT THEOREMS

① Markov Inequality - If X is a rv that takes non negative value then for any $a > 0$

$$\boxed{P[X \geq a] \leq \frac{E(X)}{a}}$$

② Chebyshev's Inequality - If X is a rv with μ & var σ^2 then for any values of $k > 0$

$$\boxed{P[|X - \mu| \geq k] \leq \frac{\sigma^2}{k^2}}$$

If a rv has small var then the probability that it takes values far away from mean is small.

→ Laws Of Large Numbers

Weak law - how a seq of prob converges.

Strong law - how a seq of rv behave in the limit

→ Central Limit Theorem (CLT)

Sum of large number of iid rv's have a distributⁿ that is approximately normal.