PROBABILITY
and
RANDOM Processes

* Sample Space: (Ω) collection of all possible outcomes (exhaustive)
* outcomes: elements in Ω (all outcomes are disjoint)
* events: subsets of Ω, if $|\Omega| = n \Rightarrow 2^n$ events possible
* Probability space: subset of power set of Ω
* Prob. Law: distribution of probabilities to Ω (follows some Rules).

1.) $P(\omega) >= 0 \;\forall\; \omega \in \Omega$

2.) $\sum_{\omega \in \Omega} P(\omega) = 1$

3.). if $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$

Discrete Modles elements of Ω are discrete

Let A be an event $A = \{s_1, s_2 \dots s_n\}$ $\quad s_i \in \Omega$

then $P(A) = \sum_{i=1}^{n} P(s_i)$ [Since $s_1, s_2 \dots s_n$ are disjoint]

↳ Discrete Probability Law

Red
Blue
Green
Black

* If $A \subset B$ then $P(A) \leq P(B)$

$$B = B \cap \Omega$$
$$= B \cap (A \cup A^c)$$
$$= (B \cap A) \cup (B \cap A^c)$$
$$B = A \cup (B \cap A^c)$$ (as A and $B \cap A^c$ are disjoint).

$$\Rightarrow P(B) = P(A) + P(B \cap A^c)$$

$$\Rightarrow \boxed{P(B) \geq P(A)}$$

* also we know

$A \cup B = A \cup (A^c \cap B)$ (disjoint)

$P[A \cup B] = P[A] + P[A^c \cap B]$

$P[B] = P[A] + P[B \cap A^c] \Rightarrow P[B] = P[A \cap B] + P[B \cap A^c]$

$P[A \cup B] - P[B] = P[A] - P[B \cap A]$

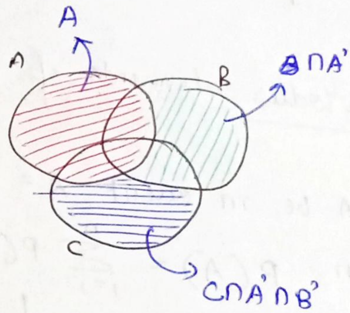
$$\boxed{P[A \cup B] = P(A) + P(B) - P(B \cap A)}$$

$$\boxed{P(A \cup B) \leq P(A) + P(B)}$$

or in general $P(\bigcup_n A_n) \leq \sum_{i=1}^{n} P(A_i)$

* $P(A \cup B \cup C)$

$= P(A) + P(B \cap A^c) + P(C \cap A^c \cap B^c)$



Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

$A_1 \cap A_2 = \phi \Rightarrow P(A_1 \cup A_2 | B) = \dfrac{P((A_1 \cup A_2) \cap B)}{P(B)}$

$= \dfrac{P(A_1 \cap B \cup A_2 \cap B)}{P(B)}$

$= \dfrac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)}$

$= \dfrac{P(A_1 \cap B)}{P(B)} + \dfrac{P(A_2 \cap B)}{P(B)}$

$= P(A_1|B) + P(A_2|B)$

#2 $P(A|B) = \frac{P(A \cap B)}{P(B)} >= 0$ and finitely defined only when $P(B) != 0$

#3. $P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

Note: since Conditional Probability satisfies all three axioms, it can be treated as a whole new prob. Law. (It's just the same prob. law with reduced sample space)

$$P(A|B) = \frac{\# \text{ in } A \cap B}{\# \text{ in } B}$$ for equally likely events.

as in an equally likely distribution, let E be an event

then $P(E) = \frac{\# \text{ in } E}{\# \text{ in } \Omega}$ or $\propto \# \text{ in } E$

*

If $A = A_1 \cap A_2 \cap \dots A_n$

$$P(A) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \dots P(A_i | \bigcap_{j=1}^{i-1} A_j) \dots P(A_n | A_1 \cap A_2 \dots A_n)$$

Multiplication Law

Def: $\bar{A}_k = A_1 \cap A_2 \cap \dots A_k$.

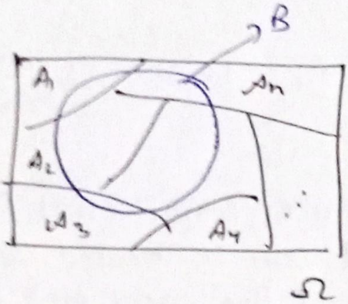
$$P(\bar{A}_n) = \frac{P(\bar{A}_n)}{P(\bar{A}_{n-1})} \cdot \frac{P(\bar{A}_{n-1})}{P(\bar{A}_{n-2})} \cdot \frac{P(\bar{A}_{n-2})}{P(\bar{A}_{n-3})} \cdots \frac{P(\bar{A}_2)}{P(\bar{A}_1)} P(\bar{A}_1)$$

E_i = no hearts is drawn in i^{th} d 13

$$P(E_1 \cap E_2 \cap E_3) = \cancel{\left(\frac{3}{4}\right) \cdot (}$$

$$\left(\frac{39}{52}\right)\left(\frac{38}{51}\right)\left(\frac{37}{50}\right)$$

Let $\Omega = \bigcup_{i=1}^{n} A_i$ (i.e. Ω is divided into n disjoint sets)



Now $B = B \cap \Omega$

but $\Omega = \bigcup_{i=1}^{n} A_i$

$\Rightarrow B = B \cap (A_1 \cup A_2 \cup \dots A_n)$

$= (B \cap A_1) \cup (B \cap A_2) \cup \dots (B \cap A_n)$

$B = \bigcup_{i=1}^{n} (B \cap A_i)$

and since A_i are disjoint $B \cap A_i$ are also disjoint

$\{ (B \cap A_i) \cap (B \cap A_j) = B \cap A_i \cap A_j \cap B = B \cap \phi = \phi \}$.

Thus $\boxed{Pr(B) = \sum_{i=1}^{n} Pr(B \cap A_i)}$

also $Pr(B \cap A_i) = Pr(B|A_i) \cdot Pr(A_i)$.

$\Rightarrow \boxed{Pr(B) = \sum_{i=1}^{n} Pr(A_i) \cdot Pr(B|A_i)}$ (Total Probability Theorem)

Q) die can be rolled at max 2 times
2nd time rolled only if ~~#~~ {1, 2} occurs in 1st roll.

B: sum >= 4

A_i: 1st die shows $i = \frac{1}{4} \; \forall i$

$\cancel{P(B) = \frac{3}{4} \frac{1}{4} \cdot 1 + \frac{2}{4}}$

A_1 : $P(B|A_1) = \frac{2}{4}$

A_2 : $P(B|A_2) = \frac{3}{4}$

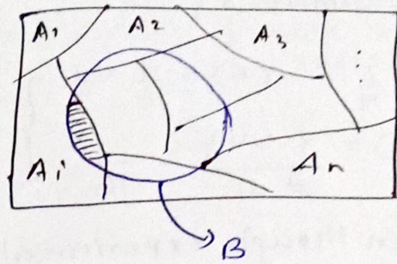
A_3 : $P(B|A_3) = 0$

A_4 : $P(B|A_4) = 1$

$P(B) = \frac{1}{4}\left(\frac{2}{4} + \frac{3}{4} + 0 + 1\right) = \frac{9}{16}$

* ## Bayes Theorem

we want to calculate



$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)}$$

but we know

$$P(A_i \cap B) = \cancel{P(B).P(}$$
$$P(B|A_i).P(A_i)$$

and $P(B) = \sum_{j=1}^{n} P(B|A_j).P(A_j)$. (from Total Probability Thm)

$$\Rightarrow \boxed{P(A_i|B) = \frac{P(B|A_i).P(A_i)}{\sum_{j=1}^{n} P(B|A_j)\,P(A_j)}}$$ (Bayes Rule)

Ex: A: Aircraft is Present
B: Radar detects Aircraft.

$P(A) = 0.05$
$P(B|A) = 0.99$
$P(B|A^c) = 0.1$

we want to find $P(A|B) = \frac{P(B|A).P(A)}{P(B|A).P(A) + P(B|A^c).P(A^c)}$

$$P(A|B) = \frac{0.99.\cancel{0.05}}{0.99.\cancel{0.05} + 0.1 \times \underset{19}{\cancel{0.95}}} = \frac{0.99}{0.99 + 1.9} = \frac{0.99}{2.89}$$

$$\begin{array}{r} 1.90 \\ 0.99 \\ \hline 2.89 \end{array}$$

$= 0.342$

* ## Independent Events

~~P(A~~ for independent events (A, B) clearly

$\boxed{P(A|B) = P(A)}$ — (I). ($P(B)$ is conditioned $!= 0$).

$\Rightarrow \boxed{P(A \cap B) = P(A).P(B)}$ — (II) (No condition: General def)

** Note: 2 mutually exclusive events are always dependent

Ex A: { 1st roll is 1 } are A, B in...

B: { sum is 5 }

$P(A) = \frac{1}{4}$ $P(B) = \frac{4}{16} = \frac{1}{4}$

$P(A \cap B) = \frac{\{1,4\}}{\# \Omega} = \frac{1}{16}$

$\Rightarrow P(A \cap B) = \frac{1}{16} = P(A) \cdot P(B) = \frac{1}{4} \cdot \frac{1}{4}$

Note: Even though experiments seem dependent intuitively, we know only when we check mathematically if they are indep./dep.

Another way to chk: check $P(A|B) = A$ and $P(B|A) = B$ intuitively for dependence.

Q). A: (max is 2)
B: (min is 2)

$P(A) = \frac{3}{16}$ $P(B) = \frac{5}{16}$ $P(A \cap B) = \frac{1}{16}$

A and B are not ~~dep~~ independent.

If 2 events are independent, the ~~must be~~ might/might not be independent even if some other event has occured

i.e. $\boxed{P(A \cap B | C) = P(A|C) \cdot P(B|C)}$ $\Rightarrow$ A, B are independent

also $P(A | B \cap C) = P(A|C)$ — Conditional ind.

$A \cap B | C$
$= \frac{P(A \cap B \cap C)}{P(C)} = \frac{P(A \cap C) P(B \cap C)}{4}$

$A \cap C$ $B \cap C$

$\boxed{P(A \cap B) = P(A) P(B)}$ — Marginal ind.

Note:

1). Ind. of A and B does not imply that A and B are cond. ind.

A: outcome of 1st die
B: outcome of 2nd die
C: sum of 2 die = 2

2) Cond. Ind. of A and B does not imply that A and B are Ind.

* independence of 3 events $A_1\, A_2\, A_3$

$\Rightarrow$ cond: $\Rightarrow P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$

$\Rightarrow$ cond $\Rightarrow P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$] both cond. need to be satisfied.

Eg:

H_i : { ith toss is Head }

D : { outcomes are diff }

H_1 H_2 and D are pairwise Independent

but $P(H_1 \cap H_2 \cap D) \neq P(H_1) \cdot P(H_2) \cdot P(D)$.

~~P(H~~ * H_1 and H_2 are ind.

* $P(H_i \cap D) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = P(H_i) \cdot P(D) = \frac{1}{2} \cdot \frac{1}{2}$

but $P(H_1 \cap H_2 \cap D) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$

Eg: Roll a 6-sided die.

A : { 1st shows {1, 2, 3} } $\Rightarrow P(A) = 3/6 = 1/2$

B : { 1st shows {3, 4, 5} } $\Rightarrow P(B) = 3/6 = 1/2$

C : { sum is 9 } $= P(C) = 1/9$

$P(A \cap B \cap C) = 1/36 = P(A) \cdot P(B) \cdot P(C)$.

$\Rightarrow$ generalizing: $A_1, A_2, \dots A_n$ are independent iff.

$$P\left(\bigcap_{i \in S} A_i\right) = \prod_{i \in S} P(A_i). \quad \{A_i\}\ i \in S \quad S \subseteq \{1, 2, \dots n\}$$

Independent Trials:

a series of independent identical experiments conducted multiple times

Eg: Coin is tossed n-times.

A_i = { ith coin toss is Head }.

$P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$.

Bernoulli Trials

$\Rightarrow$ independent trial with binary outcomes

$P(A \cap B) \geq P(A) + P(B) - 1$

$A \cup B = A + B - A \cap B$

$A \cap B = A + B - A \cup B$

$A \cap B \geq A + B - 1$

$P(A_1 \cup A_2 \cup \dots A_n) = \cancel{P(A_1) + P(A_2) + \dots}$
$\Sigma P A_i - A_1 A_2 + A_4 A_2 A_3 - \dots (-1)^{n+1} \bigcap A_i$ all

$1 - P(\overline{\overline{A_1} \cup \overline{A_2} \cup \dots \overline{A_n}}) \geq P(A_1) + P(A_2) + \dots - (n-1)$

$\Rightarrow 1 - P(\bigcup \overline{A_i}) \geq \Sigma P(A_i) - (n-1)$

$P(\bigcup B_i) \leq \Sigma P(B_i)$

$\cancel{1 = P(\bigcup \overline{A_i}) =}$ $1 - P((\bigcap A_i)^c)$

$= 1 - P(\bigcup A_i^c)$

$\geq 1 - \Sigma(1 - P(A_i))$

$\geq 1 - n - P(A_1) - P(A_2) - \dots P(A_n)$

2 - 1 + 3 - 2 +

$(A_1^c \cap A_2) \cup (A_2^c \cap A_3) \cup A_3^c \cap A_4 \dots$

$A_1, A_2 \dots \quad A_n \subseteq A_{n+1}$

$$P(A) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\lim_{n\to\infty} \bigcup_{i=1}^{n} A_i\right).$$

$$P(A) = \lim_{n\to\infty} P(A_n).$$

$B_1 = A_1$

$B_k = A_k \cap A_{k-1}^c$

$$\Rightarrow \bigcup_{k=1}^{n} B_k = \bigcup_{k=1}^{n} (A_k \cap A_{k-1}^c) = \bigcup_{k=1}^{n} A_k = A_n$$

$$\Rightarrow P\left(\bigcup_{k=1}^{n} B_k\right) = A_n = P(A).$$

as $A = \bigcup_{i=1}^{\infty} B_i$

~~P(A~~

$$P(A) = \sum_{i=1}^{\infty} P(B_i)$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} B_i\right)$$

$$P(A). = \lim_{n\to\infty} P(A_n).$$

using this Result

* $A_1, A_2 \dots$ where $A_n > A_{n+1}$ (A_n contains A_{n+1}).

$$A = \bigcap_{i=1}^{\infty} A_i$$

$$C = A^c = \bigcup_{i=1}^{\infty} A_i^c \to C_i$$

$$P(A) = 1 - P(C) = 1 - \lim_{n\to\infty} P(C_n)$$

$$P(A) = \lim_{n\to\infty} (1 - P(A_n^c))$$

$$P(A) = \lim_{n\to\infty} P(A_n)$$

Random Variables

* D.R.V: If a R.V. takes values from a finite or countably infinite set, then it is called D.R.V.

  Ex: 1) $x \in \{1, 2, 3, \ldots N\}$

  2) $x \in \mathbb{Z}$

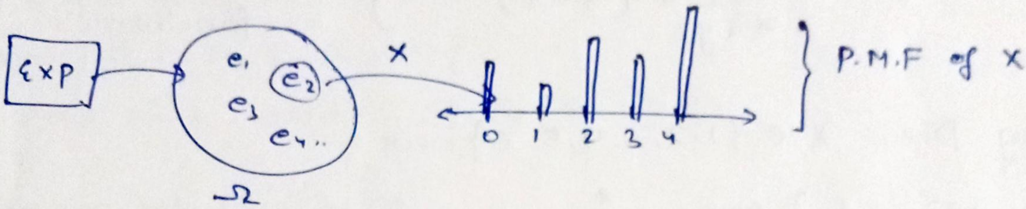
  3) $x \in \mathbb{N}$

* C.R.V: R.V. that is not discrete, is continuous R.V.

  Eg 1) $\mathbb{R}$

  2) $[0, 1]$

  3) $\{[a_i, b_i]\}_{i=1}^{N}$



then function $y = P_X(X = x)$ is called P.M.F (Prob. mass function)

* continuous analogue is PDF (Prob. distribution function).

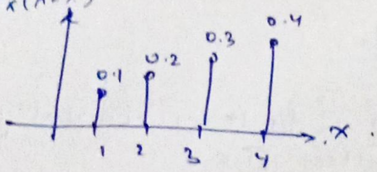
Let $A \subseteq \Omega$ then $X(\omega) = x \;\; \forall \omega \in A$ (assume)

then $\mathbb{P}_X(A) = \mathbb{P}(\{X = x\})$.

Note that $\{X = x_1\} \; \{X = x_2\} \; \{X = x_3\} \ldots$ are all disjoint.

in Prob. axiom:	in R.V
* $P(\Omega) = 1$	$\mathbb{P}\left(\bigcup_{X=x} \{X = x\}\right) = 1$
$P(\cup A_i) = 1$	$\Rightarrow \boxed{\sum_x \mathbb{P}_X(\{X = x\}) = 1}$
A_i are elements in Ω	

Eg:



* Let $S = \{2,3\}$

then $\mathbb{P}(X \in S) = \sum_{k \in S} \mathbb{P}_X(X=k)$

$= 0.2 + 0.3 = 0.5$

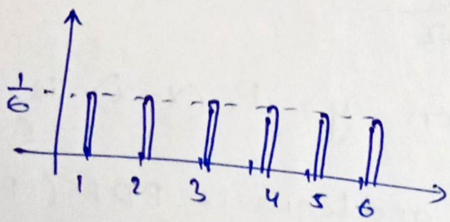
* $\cancel{\mathbb{P}(\{X>x\})} = \sum_{\cancel{y=x+1},\ y>x} \cancel{\mathbb{P}_X(X=y)}$.

$$\left.\begin{array}{l} * \ \mathbb{P}(\{X>y\}) = \sum_{x>y} \mathbb{P}_X(X=x) \end{array}\right\} \text{CCDF (Complimentary CDF)}$$

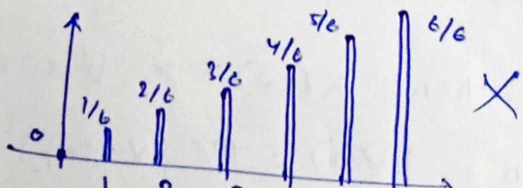
$$\left.\begin{array}{l} * \ \mathbb{P}(\{X \le y\}) = 1 - \sum_{x>y} \mathbb{P}_X(X=x) \\ \qquad = \sum_{x \le y} \mathbb{P}(X=x) \end{array}\right\} \text{CDF (Cumulative distribution function)}$$

Ex: Rolling Die. $X \in \{1,2,3,4,5,6\}$.

* $\left.\Pr_X(\{X=k\}) = \frac{1}{6}\right\}$ PMF



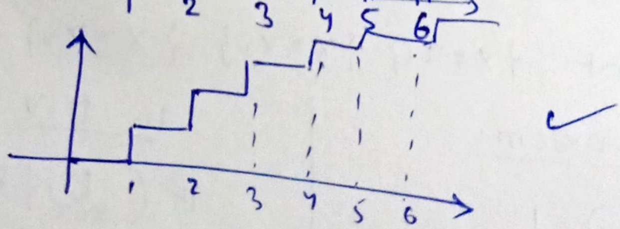
* $\Pr_X(X <= k) \Rightarrow$ CDF.



Properties:

1) monotonically non decreasing

2) Right Continuous.



3) $\Pr_X(X <= \infty) = 1$.

$\Pr_X(X <= -\infty) = 0$

$X = \begin{cases} 1 & \text{if success with prob } p \\ 0 & \text{if failure with prob } 1-p \end{cases}$

$X = \{0, 1\}$. $\Omega = \{S, F\}$. $P_X(0) = 1-p$
$P_X(1) = p$.

for prob. space to be valid,

$\sum_{x \in X} P_X(x) = 1 \Rightarrow P_X(0) + P_X(1) = 1-p+p = 1$
$\Rightarrow$ Valid PMF.

2) Binomial R.V (n, p).

counts number of successes in n trials (independent).

$X = \{\#$ no. of successes in 'n' ind. trials$\}$

$X = \{0, 1, 2, 3, \ldots, n\}$

for PMF to be valid. $\sum_{x \in X} P_X(X = x) = 1$.

defining Prob. Law for X.

$P_X(X = k) = {}^nC_k\, p^k (1-p)^{n-k}$

checking validity of PMF of X

$\sum_{x \in X} P_X(X = x) = 1 \Rightarrow \sum_{x=0}^{n} {}^nC_x\, p^x (1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1$
(using Bino. expansion).

3) Geometric. R.V (p)

$X := \{\#$ trials after which 1st success occurs$\}$.

$X = \{1, 2, \ldots \infty\} = \mathbb{N}$

$P_X(X = k) = (1-p)^{k-1} \cdot p$.

checking validity.

$\sum_{k \in X} P_X(X = k) = 1 \Rightarrow \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p = p[1 + (1-p) + (1-p)^2 + \ldots]$

$= \frac{p}{1-(1-p)} = \frac{p}{p} = 1$

Poisson

4) ~~Gaussian~~ R.V ($\lambda = np$)

Consider a Binomial R.V where $n \to \infty$, $p \to 0$ such that $\lambda = np = \text{const.}$

$$\mathbb{P}_X(X=k) = {}^nC_k (p)^k (1-p)^{n-k}$$

$$= \frac{n(n-1)\cdots(n-k+1)}{n^k \cdot k!} (np)^k \left(1 - \frac{np}{n}\right)^{n-k}$$

$$= \left\{\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-k+1}{n}\right\} \left(\frac{\lambda}{k!}\right)^k \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k}$$

Calc for $\lim n \to \infty$

$$= \{1 \cdot 1 \cdot 1 \cdots 1\} \left(\frac{\lambda^k}{k!}\right) e^{-\lambda} \cdot 1$$

$$= e^{-\lambda} \frac{\lambda^k}{k!}$$

$$\boxed{\mathbb{P}_X(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}}$$

checking $\displaystyle\sum_{k \in X} P_X(X=k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!}$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \left\{\frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^k}{k}\right\}$$

$$= e^{-\lambda} \cdot e^{\lambda} = 1$$

Mean: (Expectation of R.V)

$$\mathbb{E}[X] = \sum_{x \in X} x \cdot \mathbb{P}_X(X=x).$$

Variance:

$$\sigma_X^2 = \text{var}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right]$$

$$\sigma_X = \sqrt{\text{var}(X)} = \text{S.D.}$$

* when $\mathbb{E}[X]=0$, σ_X represent R.M.S. value of X.

$$\sigma_X = \sqrt{\mathbb{E}[X^2]} = \text{RMS value of } X$$

* n^{th} moment of a R.V. $= \mathbb{E}[X^n]$

▷ Function of a R.V

Let $Y = g(X)$.

then $$\boxed{\mathbb{P}_Y[Y=y] = \sum_{x:\, g(x)=y} \mathbb{P}_X[X=x]}$$

$$\mathbb{E}[Y] = \sum_{y\in Y} y \cdot P_Y[y].$$

$$= \sum_{y\in Y} y\left(\sum_{x:\, g(x)=y} \mathbb{P}_X(x)\right).$$

$$= \sum_{y} \sum_{x:\, g(x)=y} g(x) \cdot \mathbb{P}_X(x).$$

Let $A_y = \{x : g(x) = y\}$.

Now $A_{y_1}, A_{y_2}, \dots$ A_{y_i} are all disjoint since a function cannot be one $\rightarrow$ many mapping.

$$\Rightarrow \left(\sum_{y} \sum_{x:\, g(x)=y} \equiv \sum_{x}\right)$$

$$\Rightarrow \boxed{\mathbb{E}[Y] = \sum_{x\in X} g(x) \cdot \mathbb{P}_X(x).}$$

$\Rightarrow$ Results:

1) $\mathrm{Var}(X) = \sum_{x} (x - \mathbb{E}(X))^2 \cdot \mathbb{P}_X(x)$.

2) $\mathbb{E}[X^n] = \sum_{x} x^n \, \mathbb{P}_{r_X}(x)$

Consider Y as a linear function of X.

* $Y = a.X + b$.

$\Rightarrow \mathbb{E}[Y] = \sum(ax+b)\, \mathbb{P}r_X(x)$

$= \sum ax.\, \mathbb{P}r_X(x) + \sum b.\, \mathbb{P}r_X(x)$

$= a. \sum_x x.P_x(x) + b.\sum_x P_x(x)$

$= a.\mathbb{E}[X] + b.1$

$\mathbb{E}[Y] = a.\mathbb{E}[X] + b$

Note that this is valid only for Y being linear function of X

* $var[Y] = \mathbb{E}((Y-\mathbb{E}[Y])^2)$

$= \sum_y (Y-\mathbb{E}[Y])^2\, P_Y(y)$.

$= \sum_{x} (a X + \cancel{b} - a\mathbb{E}[X] - \cancel{b})^2\, P_x(x)$.

$= \sum_x a^2 (X - \mathbb{E}[X])^2\, P_x(x)$

$= a^2 \sum_x (X - \mathbb{E}[X])^2\, P_x(x)$

$var(Y) = a^2 (var(X))$.

This also works only if Y = Linear function of X.

* $var[X] = \sum (X - \mathbb{E}[X])^2\, P_X(x)$

$= \sum [X^2 + \mathbb{E}[X]^2 - 2X\,\mathbb{E}[X])\, P_X(x)$

$= \sum X^2.P_X(x) + \sum \mathbb{E}[X]^2.P_X(x) - 2\sum X\,\mathbb{E}[X].P(X)$

$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$

E of Bernoullie

$X :=$ Bernoullie R.V. $(P) = \begin{cases} 1 & p \\ 0 & (1-p) \end{cases}$

$$\mathbb{E}(X) = 1 \cdot p + 0 \cdot (1-p) = p$$

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= p - p^2 \\ &= p(1-p) = pq \end{aligned}$$

$X :=$ Poisson R.V.

$$* \quad \mathbb{E}[X] = \sum_k k \cdot \frac{e^{-\lambda} \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_k \frac{\lambda^k}{(k-1)!} \frac{\lambda}{\lambda}$$

$$e^x = \frac{x^0}{0!} + \frac{x^1}{1!} + \ldots$$

$$= \frac{\lambda e^{-\lambda}}{\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

$$= \lambda e^{-\lambda} \cdot e^{\lambda}$$

$$\boxed{\mathbb{E}[X] = \lambda}$$

$$* \quad \mathbb{E}[X^2] = \sum_k k^2 \frac{e^{-\lambda} \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_k \frac{\lambda^k \cancel{k^2} (k+1-1)}{\cancel{k}(k-1)!}$$

$$= e^{-\lambda} \sum_k \frac{\lambda^k (k-1) + \lambda^k}{(k-1)!}$$

$$= e^{-\lambda} \left[\sum_{k=1}^{\infty} \frac{\lambda^2 \lambda^k}{(k-2)! \lambda^2} + \lambda \sum_k \frac{\lambda^k}{(k-1)! \lambda} \right]$$

$$\cancel{= e^{-\lambda} [e^{\lambda} + e^{\lambda}] = 2.}$$

$$= e^{-\lambda} (\lambda^2 \cdot e^{+\lambda} + \lambda \cdot e^{+\lambda})$$

$$= \lambda^2 + \lambda = \lambda(\lambda + 1).$$

$\mathbb{E}[X^2] = \lambda^2 + \lambda$

$\mathbb{E}[X] = \lambda$

$\Rightarrow \mathrm{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$

$$\Rightarrow \left.\begin{array}{l} \mathbb{E}[X] = \lambda \\ \mathrm{Var}[X] = \lambda \end{array}\right\} \text{ for Poisson } X$$

<u>Joint PMF of 2 R.V.s</u> (X and Y)

$P_{X,Y}(x, y) = P[X = x, Y = y]$.

to find marginal PMF of X from given joint PMF (X,Y)

$$\left.\boxed{P_X(X=x) = \sum_{\substack{y \in Y \\ x = x}} P_{X,Y}(x, y)}\right\} \text{ marginal PMF of } X.$$

also, we know,

$$\mathbb{E}[g(X,Y)] = \sum_x \sum_y g(x, y) . P_{X,Y}(x, y)$$

<u>Ex:</u> $\mathbb{E}[X+Y] = \sum_x \sum_y (x+y)\, P_{X,Y}(X=x, Y=y)$

$$= \sum_x \left[\sum_y x\, P_{X,Y}(x, y) + \sum_y y\, P_{X,Y}(x, y)\right]$$

$$= \sum_x \left[x . P_X(x) + \sum_y y . P_{X,Y}(x, y)\right]$$

$$= \sum_x . x\, P_X(x) + \sum_y . \sum_x y . P_{X,Y}(x, y)$$

$$= \sum_x . x . P_X(x) + \sum_y y . P_Y(y)$$

$$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

In general.

$$\boxed{\mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = \sum_{i=1}^{N} \mathbb{E}[X_i]}$$

for identical distribution $\mathbb{E}[X_i] = \mathbb{E}[X_j] \ \forall i,j$

$$\Rightarrow \quad \mathbb{E}\left[\sum_{i=1}^{N} X_i\right] = N.\mathbb{E}[X_1]$$

▶ Conditional PMF (conditional on an event A).

$$P_{X|A}(x) = \frac{P(\{X=x\} \cap A)}{P(A)}.$$

$$P(A) = \sum_{x \in A} P_X(x).$$

$$\sum P_{X|A}(x) = \sum_{x \in A} \frac{P(X=x)}{P(A)} = 1$$

Eg: $\cancel{P_{X|A} = P}$ $K \sim$ Geometric R.V.

$A = \{$student passes test in atmost N attempts$\}$.

$$P_{K|A} = \frac{P_X(\{K=k \cap A\})}{P_X(A)} = \frac{(1-p)^{k-1} \cdot p'}{\sum_{m=1}^{N} (1-p)^{m-1} \cdot p'}$$

* conditional on another R.V (Y)

$$P_{X|Y}(x) = \frac{P_{X,Y}(x,y)}{P_Y(y)} \quad \text{or}$$

$$P_{X,Y}(x,y) = P_{X|Y}(x).P_Y(y)$$
$$= P_{Y|X}(y).P_X(x).$$

if X and Y are independent, then

$$\boxed{P_{X,Y}(x,y) = P_X(x).\ P_Y(y)}$$

$\mathbb{E}[X|Y]$ is itself a R.V

$$\mathbb{E}[\mathbb{E}[X|Y]] = \sum_y \left(\sum_x x \cdot P_{X|Y}(x) \right) P_Y(y)$$

$$= \sum_y \sum_x x \cdot P_{X,Y}(x,y)$$

$$= \sum_x x \cdot P_X(x)$$

$$= \mathbb{E}[X]$$

Thus, $\boxed{\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X]}$

we used the result that

$$\mathbb{E}[X|Y] = \sum_{x \in X} x \cdot P_{X|Y}(x|\cancel{Y}y)$$

also, note that

$\mathbb{E}[X|Y] = \mathbb{E}[X]$ when X, Y are independent.

* if X and Y are independent,

① $\mathbb{E}[X \cdot Y] = ?$

$$\mathbb{E}[X \cdot Y] = \sum_X \sum_Y (x \cdot y) \, \mathbb{P}_{X,Y}(X = x, Y = y)$$

$$= \sum_X \sum_Y x \cdot y \; \mathbb{P}_X(x) \cdot \mathbb{P}_Y(y)$$

$$= \sum_y \mathbb{E}[X] \cdot y \cdot P_Y(y)$$

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

② $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ (regardless of their independence).

$Var(X+Y) = ?$

$$Var(X+Y) = \mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2$$

$$= \mathbb{E}[X^2 + Y^2 + 2X.Y] - \mathbb{E}[X+Y]^2$$

$$= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + \mathbb{E}[2.X.Y] - (E(X)^2 + \mathbb{E}(Y)^2 + 2E(X).E(Y))$$

$$= \sum_x x^2 P(x) + \sum_y y^2 P(y) + 2E[X.Y] - E(X)^2 - E(Y)^2 - 2E(X).E(Y)$$

$$= var(X) + Var(Y) + 2\mathbb{E}[X.Y] - 2E(X).E(Y)$$

$$= Var(X) + Var(Y) + 2(\mathbb{E}[X.Y] - \mathbb{E}[X].\mathbb{E}[Y])$$

but $\mathbb{E}[X.Y] = \mathbb{E}[X].\mathbb{E}[Y]$ as. X, Y are ind.

$$\Rightarrow \boxed{Var(X+Y) = Var(X) + Var(Y)}$$

** if X and Y are independent.

in general.

$$var(X+Y) = var(X) + var(Y) + 2\underbrace{(\mathbb{E}[X.Y] - \mathbb{E}[X].\mathbb{E}[Y])}_{\substack{cov(X,Y) \\ (covariance)}}.$$

$$Cov(X,Y) = \mathbb{E}[XY - E[X]E[Y]].$$

Ex: find $\mathbb{E}$ and Var of Bino X.

$$P_X(x) = {}^nC_x \, p^x (1-p)^{n-x}.$$

$$\Rightarrow \mathbb{E}[X] = \sum_{x=0}^{n} x.\, {}^nC_x \, p^x (1-p)^x$$

we can say that 1 Binomial R.V. is just N. independent identical distributed (iid). Bernoullie. R.V.

$$\Rightarrow E[X_1 + X_2 .. X_n] = E[X_1] + E[X_2] + .. + E[X_n].$$

$\Rightarrow X_i$ are iid $\qquad Y = NX_i$

$$\mathbb{E}[N.X_i] = N.\mathbb{E}[X_i]$$

$$\mathbb{E}[Y]. = N.p$$

sim. from prev result. $Var(Y) = N(pq)$

Result:

$$E[\mathrm{var}(Y|X)] = E[E[Y^2|X]] - E[E[Y|X]^2)$$

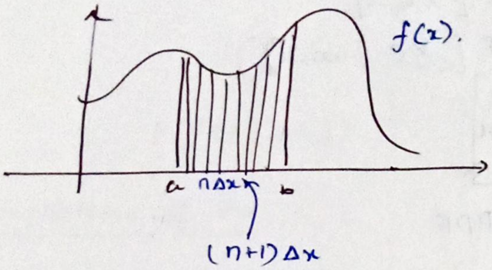
$$= E[Y^2] - E[E[Y|X]^2)$$

$$\mathrm{Var}[E[Y|X]) = E[E[Y|X]^2) - (E[E[Y|X]])^2$$

$$\mathrm{va}\cancel{\mathrm{r}[Y]} =$$

$$\mathrm{var}[Y) = E[\mathrm{var}[Y|X]) + \mathrm{Var}(E[Y|X])$$

Continuous Random Variable



Let $A = [a, b]$.

$$[a, b] = \bigcup_n [n\Delta x, (n+1)\Delta x]$$

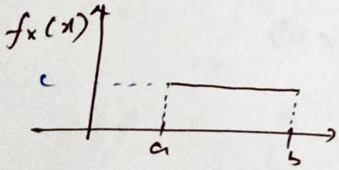
$$P(A) = P(X \in [a,b]) = \sum_{n=1}^{N} P(X \in (n\Delta x, (n+1)\Delta x))$$

$$= \sum_{n=1}^{N} f_X(n\Delta x).\Delta x \quad \text{(This is how we define it)}$$

$$P(A) = \int_a^b f_X(x).\,dx$$

$$\boxed{P(A) = \int_a^b f_X(x)\,dx}$$ Note that $$\boxed{P(X=k) = \int_k^k f_X(x)\,dx = 0}$$

Eg: Let $X \sim U(a,b)$ [X is uniformly distributed in (a,b)].



$$P(\Omega) = P[X \in [-\infty, \infty]]$$

$$P(\Omega) = \int_{-\infty}^{\infty} f_X(x)\,dx = 1. \quad \text{should satisfy.}$$

$\Rightarrow$ finding c, $$\int_a^b c.\,dx = (b-a).c = 1$$

$$\Rightarrow \boxed{c = \frac{1}{b-a}}$$

thus. $$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in (a,b) \\ 0 & \text{otherwise} \end{cases}$$

Note that when $b = a$, this function $f_X(x)$ becomes ~~f(x)~~. $\delta(x-a)$.

CDF

we define a function $F_X(x) = P[X \leq x]$

$= P[X \in [-\infty, x])$

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx}$$

CDF PDF

conversly, $f_X(x) = \frac{d}{dx}(F_X(x))$

Note that in discrete case

~~$F_X(x) = \sum_{k \leq x} P_X(k)$~~. $F_X(k) = \sum_{m \leq k} P_X(m)$.

also $\boxed{P_X(k) = F_X(k) - F_X(k-1)}$

Example: define $X = \max\{X_1, X_2, \ldots X_n\}$.

$\Rightarrow F_X(x) = P[X_1 < x, X_2 < x, \ldots X_n < x]$. if X_i are independent

$= P[X_1 < x] \cdot P(X_2 < x) \ldots P(X_n < x)$

$= \prod_{i=1}^{N} P[X_i < x]$.

$$F_X(x) = \prod_{i=1}^{N} F_{X_i}(x)$$

when X_i are iid, $F_{X_i}(x)$ are all same

$$\Rightarrow F_X(x) = (F_{X_1}(x))^N$$

* Properties of CDF

1) Non decreasing $F_X(x+\delta) \geq F_X(x)$.

2). $F_X(\infty) = 1$

$F_X(-\infty) = 0$

3) $P(X \in (a, b]) = F_X(b) - F_X(a)$.

since $P_X(X \in (a,b]) = \int_a^b f_X(x)\,dx$

$$= \int_{-\infty}^{b} f_X(x)\,dx - \int_{-\infty}^{a} f_X(x)\,dx$$

$$P_X(X \in (a,b]) = F_X(b) - F_X(a).$$

Expectation of R.V

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx$$

Variance of R.V

$$Var(X) = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2\,dx$$

Mean of g(X)

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f_X(x)\,dx$$

Eg: $x \sim U[(a,b)]$

$$E[X] = \int_{-\infty}^{\infty} x \cdot \frac{1}{(b-a)}\,dx = \frac{[x^2]_a^b}{2(b-a} = \frac{b+a}{2}$$

$$\mathbb{E}[X^2] = \frac{[x^3)}{3\ (b-a)} = \frac{a^2+b^2+ab}{3}$$

$$var(X) = E[X^2] - E[X]^2 = \frac{4a^2+4b^2+4ab-3a^2-3b^2-6ab}{12}$$

$$var(X). = \frac{a^2+b^2-2ab}{12} = \frac{(a-b)^2}{12}$$

Eg: $X \sim Exp.(\lambda)$.

Let $f_X(x) = \lambda \cdot e^{-\lambda x}$ for $x > 0$.

$$\int_0^{\infty} f_X(x). = \frac{\lambda}{(-\lambda)}.[0-1] = 1$$

$$\mathbb{E}[X] = \int_{0}^{\infty} x \cdot \lambda e^{-\lambda x} = \lambda \int_0^{\infty} x \cdot e^{-\lambda x}\,dx = \cancel{\int_0^{\infty}} \quad \cancel{dt}$$

$\lambda x \cancel{= t \Rightarrow \lambda dx = dt}$

$$E[X] = \lambda \int_0^\infty x\, e^{-\lambda x} dx = \lambda^{-1}$$

$$E[X^2] = \lambda \int x^2 e^{-\lambda x} dx$$

$$= \lambda \left[\frac{x^2 e^{-\lambda x}}{(-\lambda)} + \frac{1}{\lambda} \int 2x \cdot e^{-\lambda x} \right]$$

$$= \lambda \left[\left(\frac{x^2 e^{-\lambda x}}{-\lambda} \right)_0^\infty + \frac{2}{\lambda^2} \cdot \frac{1}{\lambda} \right]$$

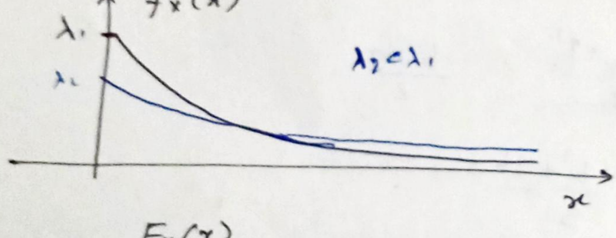
$$E[X^2] = \lambda \left(0 - 0 + \frac{2}{\lambda^3} \right) = \frac{2}{\lambda^2}$$

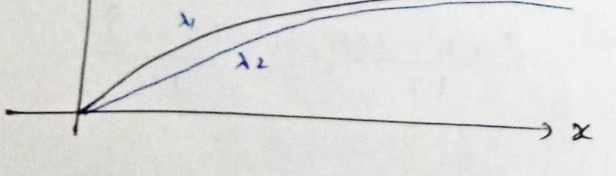
$$\boxed{Var(X) = \frac{1}{\lambda^2}}$$

$$* \quad F_X(x) = \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x}$$

$$1 - \frac{1}{e^{\lambda x}}$$

Drawing graphs





$\lambda_2 < \lambda_1$

$$\cancel{e^{-(\lambda_1 + \lambda_2)x}}$$

$$\frac{d}{dx} \left(1 - e^{-\lambda_1 x}\right)\left(1 - e^{-\lambda_2 x}\right)$$

$$1 - e^{-\lambda_1 x} - e^{-\lambda_2 x} - e^{-(\lambda_1 + \lambda_2)x}$$

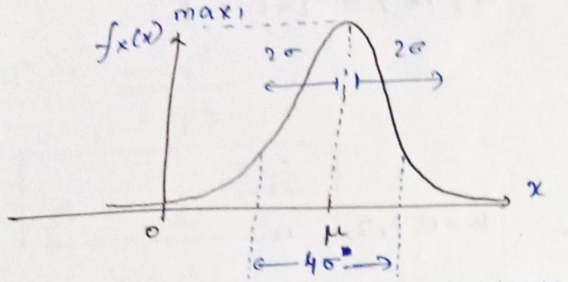
$$0 \quad \lambda_1 e^{-\lambda_1 x} + \lambda_2 e^{-\lambda_2 x} + (\lambda_1 + \lambda_2) e^{-(\lambda_1 + \lambda_2)x}$$

Normal / Guassian Dist.

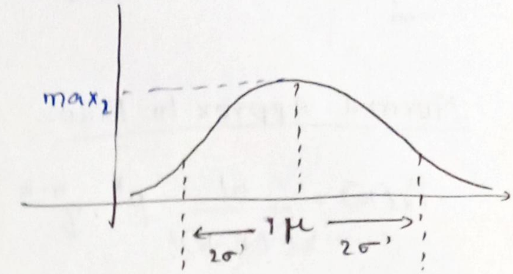
$X \sim N(\mu, \sigma^2)$

μ ← mean, σ^2 ← variance

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$$ for $x \in \mathbb{R}$



(~95% of the distribution)

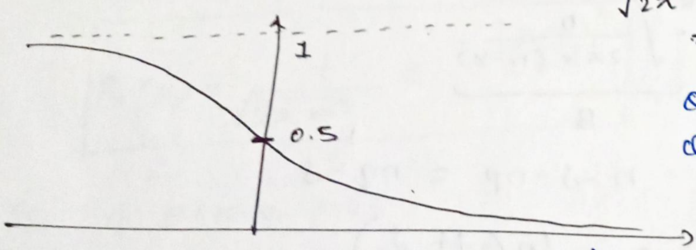


$\sigma' > \sigma$ and $max_2 < max_1$

Finding CDF

$F_X(x) = \int_{-\infty}^{x} f_X(x)$. Let us define Q function

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp\left(-\frac{u^2}{2}\right) du$$



$Q(-\infty) = 1$

$Q(\infty) = 0$

$Q(-x) = 1 - Q(x)$.

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\,\sigma}$$ let $\frac{x-\mu}{\sigma} = t$, $\frac{dx}{\sigma} = dt$

$$F_X(x) = \int_{-\infty}^{x-\mu} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x-\mu} e^{-t^2/2} dt \Rightarrow \cancel{F_X(x) = 1 - Q(}$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-y^2/2)\, dy - \frac{1}{\sqrt{2\pi}} \int_{\frac{x-\mu}{\sigma}}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy$$

$$\boxed{F_X(x) = 1 - Q\left(\frac{x-\mu}{\sigma}\right)}$$

$$\operatorname{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right].$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = F_X(\infty) = 1 - Q(\infty) = 1 - 0 = 1.$$

Normal Approx to Bino.

$$P_X(k) = \frac{n!}{k!\,(n-k)!}\, p^k \cdot q^{n-k}, \quad k = 0, 1, 2, \dots n$$

Stirling's formula: $\boxed{n! = n^n \exp(-n)\sqrt{2\pi n}}$ (approximation)

$$\Rightarrow P_X(k) = \frac{n^n \cancel{e^{-n}} \sqrt{2\pi n}}{k^k \cancel{e^{-k}} \sqrt{2\pi k}\,(n-k)^{n-k} \cancel{e^{-(n-k)}} \sqrt{2\pi(n-k)}} \cdot p^k q^{n-k}$$

$$= \underbrace{\left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k}}_{A} \cdot \underbrace{\sqrt{\frac{n}{2\pi k\,(n-k)}}}_{B}$$

Let $k = \delta + np \Rightarrow n-k = n - \delta - np = nq - \delta$

$$\ln\left(\frac{np}{k}\right) = \ln\left(\frac{np}{\delta + np}\right) = -\ln\left(1 + \frac{\delta}{np}\right)$$

$$\ln\left(\frac{nq}{n-k}\right) = -\ln\left(1 - \frac{\delta}{nq}\right)$$

$$\ln(A) = -k\,\ln\left(1 + \frac{\delta}{np}\right) - (n-k)\ln\left(1 - \frac{\delta}{nq}\right).$$

$$= -(\delta + np)\left[\frac{\delta}{np} - \frac{1}{2}\frac{\delta^2}{(np)^2}\right] + (\delta - nq)\left[\frac{-\delta}{nq} - \frac{1}{2}\frac{\delta^2}{(nq)^2}\right]$$

(Taking first order approx of $\ln(1+x)$)

as $n \to \infty$ (ignoring $\frac{1}{n^2}$ terms and keeping $\frac{1}{n}$ terms)

~~$P_x(k) =$~~

$$\ln(A) = -\frac{\delta^2}{np} - \cancel{\delta} + \frac{1}{2}\frac{\delta^2}{np} - \frac{\delta^2}{nq} + \cancel{\delta} + \frac{1}{2}\frac{\delta^2}{nq}$$

$$\ln(A) = -\frac{1}{2}\frac{\delta^2}{np} - \frac{1}{2}\frac{\delta^2}{nq}$$

$$\ln(A) = -\frac{1}{2}\frac{\delta^2}{n}\left(\frac{1}{p} + \frac{1}{q}\right)$$

$$\ln(A) = \frac{-\delta^2}{2npq}$$

$$\Rightarrow \boxed{A = e^{\frac{-\delta^2}{2npq}}}$$

$$B = \sqrt{\frac{n}{2\pi(\delta + np)(nq - \delta)}} \simeq \sqrt{\frac{1}{2\pi npq}}$$

$$\Rightarrow \boxed{P_x(k) = \frac{1}{\sqrt{2\pi npq}}\, e^{\frac{-\delta^2}{2npq}}}$$

we see that

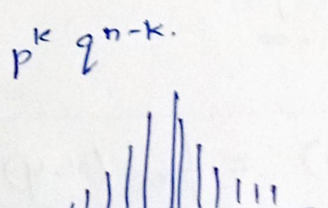
$$\boxed{\begin{aligned}\sigma^2 &= npq \\ \mu &= np\end{aligned}}$$

$$\left.\boxed{P_x(k) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(k-\mu)^2}{2\sigma^2}}}\right\}$$ Gaussian Distribution.

<u>Eg:</u> if $N = 1000$, $p = 0.5$

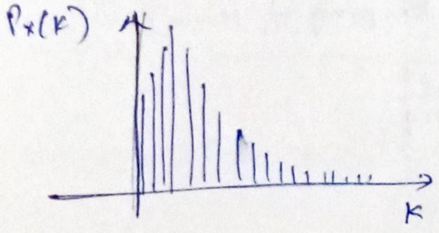
$$P(450 \leq x \leq 550) = \sum_{k=450}^{550} {}^nC_k\, p^k q^{n-k}.$$

$$x \sim N(np, npq).$$



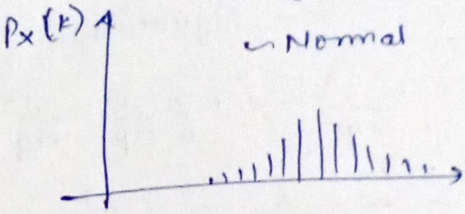
$$P(450 \leq x \leq 550) = Q\left(\frac{450 - np}{\sqrt{npq}}\right) - Q\left(\frac{550 - np}{\sqrt{npq}}\right).$$

$np = \lambda$



- Poisson (np

N large, p > 0



- Normal

- Gaussian
$N(np, \sqrt{npq})$

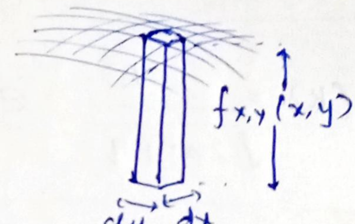
Joint PDF of X and Y

$f_{X,Y}(x,y)$.

$$P(X \in [a,b], Y \in [c,d]) = \int_c^d \int_a^b f_{XY}(x,y)\, dx\, dy.$$

~~for~~ Note: $P(x \in [a, a+dx], y \in [b, b+dy])$

$$= f_{X,Y}(x,y)\, dx.dy.$$



Joint CDF of X, Y

$$F_{X,Y}(x,y) = P[X \leq x, Y \leq y].$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\, ds.dt.$$

$$\boxed{\frac{d^2}{dx\, dy} F_{XY}(x,y) = f_{X,Y}(x,y)}$$

* Marginal Distribution from Joint distribution

$$\boxed{f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy}$$

$$P[X \in [x, x+dx]] = P([x \in (x, x+dx), y \in (-\infty, \infty)])$$

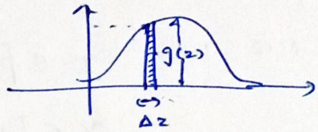
$$= f_X(x).\Delta x$$

$$= \int_{x}^{x+\Delta x} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy.dx$$

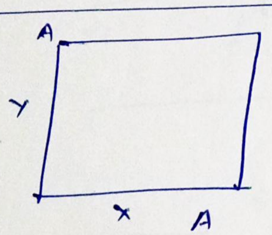
$$= \left[\int_{-\infty}^{\infty} f_{X,Y}(x,y)\right].\Delta x$$

$$\Rightarrow f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y).$$

used the fact that $\int_{z}^{z+\Delta z} g(s)\,ds = g(z)\,\Delta z$



Ex:



$f_{X,Y}(x,y) = k. \quad 0<x<A, \; 0<y<A$

and volume $= k.A^2 = 1 \Rightarrow k = \frac{1}{A^2}$

$\Rightarrow f_{X,Y}(x,y) = \frac{1}{A^2} \quad \forall x, y.$

$$F_{X,Y}(x,y) = \int_0^x \int_0^y f_{XY}(x,y)\,dx.dy = \frac{xy}{A^2}$$