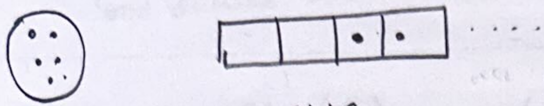
$\therefore n = 10 \ (0-9)$

$26 \times 25 \times 10 \times 9 \times 8$

Permutations: number of ways of ordering n elements.



slots

for first element $\rightarrow n$ choices

2nd element $\rightarrow (n-1)$ choices

3rd element $\rightarrow (n-2)$ choices available.

$n(n-1)(n-2)\ldots\ldots 1 = n!$

<u>Example</u>: Number of subsets of $\{1, \ldots, n\}$:

$2 \times 2 \times 2 \ldots\ldots 2 = 2^n$

$\downarrow$

put inside set or not

'Sequential way'

2^n ways

$n = 1$ $\{1\}$ $\rightarrow$ $\{1\}$ set itself, $\{\}$ null set.

$n = 2$ $\{1, 2\}$ $\rightarrow$ $\{1\}$ $\{2\}$ $\{\}$ $\{1, 2\}$

<u>Die roll example</u>

Six rolls of a (6 sided) die all give different numbers.

$6 \times 5 \times 4 \times 3 \times 2 \times 1$ [Equally likely] $\rightarrow$ Non repeating.

$$P(A) = \frac{\#\text{ in } A}{\#\text{ outcomes}} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{6^6} = 0.0154.$$

$(1, 2, 3, 4, 5, 6)$

$\downarrow$

Non repititative.

six dice = may give same o/p

$A \rightarrow$ different o/p s

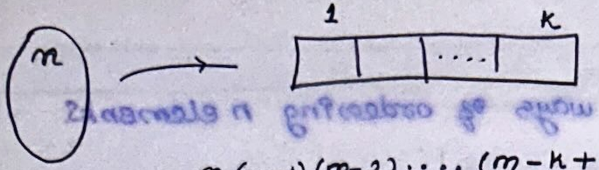
<u>Combinations</u>

n (k)

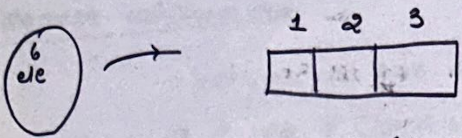
set: n elements

Definition: $\binom{n}{k} \rightarrow$ number of k-element subsets of a n-element set.

* Two ways of constructing an ordered sequence of k distinct items:



$$m(m-1)(m-2)\cdots(m-k+1) = ?$$



$$6\cdot(6-1)\cdot(6-2)$$

$$m\cdot(m-1)\cdot(m-k+1)$$

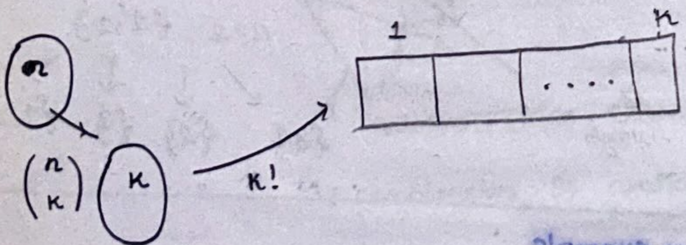
$$m(m-1)(m-2)\cdots(m-k+1)\underline{(m-k)(m-k-1)(m-k-2)\cdots 1} = m!$$

$$m(m-1)(m-2)\cdots(m-k+1)\,(m-k)! = m!$$

$$m(m-1)(m-2)\cdots(m-k+1) = \frac{m!}{(m-k)!}$$

Two ways of constructing an ordered sequence of k distinct items:

* choose the k items one at a time.
* choose k items, then order them.



① we are choosing k item subset - then arranging it in a list.

$$\text{stage 1} \times \text{stage 2} = \binom{n}{k} k!$$

$$\binom{n}{k} k! = \frac{n!}{(n-k)!}$$

$$\boxed{\binom{n}{k} = \frac{n!}{(n-k)!\,k!}}$$

$n = 0, 1, 2, \ldots$

$k = 0, 1, 2, \ldots, n$

Extreme cases

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad k = n$$

$$\binom{n}{n} = \frac{n!}{n!\,(n-n)!} = \frac{1}{0!} = \underline{1} \qquad (0! = 1)$$

ii) $$\binom{n}{0} = \frac{n!}{0!\;n!} = \frac{1}{0!} = 1$$

one subset has 0 elements — ϕ'set.

$$\sum_{k=0}^{n} \binom{n}{k} = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = \text{no. of all subsets.} = 2^n$$

$$= 1 + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \cdots + 1$$

$$= 1 + \frac{n!}{1!\,(n-1)!} + \frac{n!}{2!\,(n-2)!} + \cdots + 1$$

$$= 2 + \cdots\cdots$$

<u>we can prove using binomial theorem</u>

Binomial co.eff $\binom{n}{k}$ $\rightarrow$ Binomial probabilities.

* $n \geq 1$ independent coin tosses; $P(H) = P$, $P(T) = 1-P$

$P(k \text{ heads}) = ?$

$P(HTTHHH) = P \cdot (1-P)(1-P)\,P\,P\,P$
$n = 6$
$= P^4 (1-P)^2$

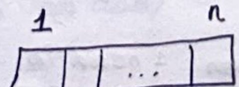
$P(\text{particular sequence}) = P^{\#\text{no. of heads}} \cdot (1-P)^{n - \#\text{ of heads}}$

$P(\text{particular } k\text{-head sequence}) = P^k (1-P)^{n-k}$

$P(k \text{ heads}) = P^k (1-P)^{n-k} \cdot (\text{Numbers of sequences.})$

$\therefore$ If no. of tosses – fixed
heads – fixed (k)
Tails – also fixed

] combinations available.
n tosses
k heads



which slot has a slot. [k slots of n has head]
$\downarrow$
Choosing k
$\downarrow$
combination of k.

$$P(k \text{ heads}) = \binom{n}{k} p^k (1-p)^{n-k}$$

[we know k subset and combinations of k in n set elements]

coin toss

* 3 heads in ten tosses
* probability that first two tosses were head.

event A: first 2 tosses were heads

event B: 3 out of 10 tosses were heads

Assume:

* Independence
* $P(H) = P$

$$P(k \text{ heads}) = \binom{n}{k} P^k (1-P)^{n-k}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(H_1 H_2 \cap 1 \text{ heads in remaining tosses})}{P(B)}$$

$$= \frac{P(H_1 H_2) \cdot P(1 \text{ heads in remaining tosses})}{P(B)}$$

$$= \left(p^2 \cdot \binom{8}{1} p^1 (1-p)^7\right) \Big/ p^3 (1-p)^7 \cdot \binom{10}{3}$$

$$= \frac{p^3 \binom{8}{1} (1-p)^7}{p^3 (1-p)^7 \binom{10}{3}}$$

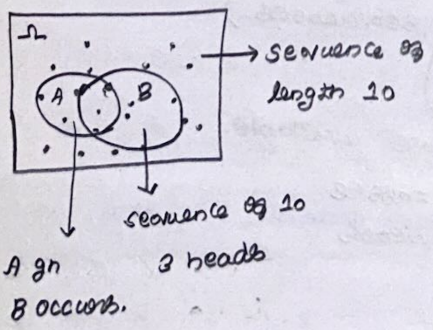
$$= \frac{\binom{8}{1}}{\binom{10}{3}} = \frac{8}{\frac{10!}{3!\,7!}} = \frac{8}{\frac{8 \times 9 \times 10}{2 \times 3}}$$

1 item out of 8 item

$$= \frac{2 \times 3}{9 \times 10} = \frac{1}{3 \times 5} = \frac{1}{15}$$

$$= 0.06667$$

Second approach



$$P(3 \text{ head}) = p^3 (1-p)^7$$ [Any 3 head sequence]

$$= \frac{\# \text{ in } A \cap B}{\# \text{ in } B} = \frac{\text{2 heads in 2 tosses then 1 head in 8 tosses}}{\binom{10}{3}}$$

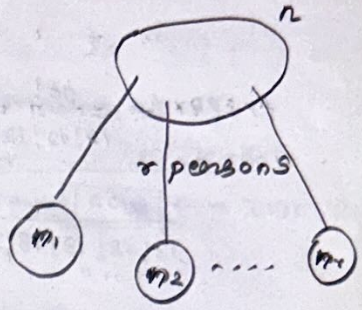
$$= \frac{8}{\binom{10}{3}}$$

Partitions

* $n \geq 1$ distinct items; $r \geq 1$ persons give n_i items to person i

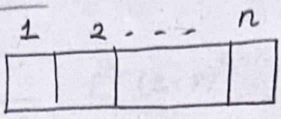
  $n_1, \ldots n_r \to$ non negative integers.

  $\boxed{n_1 + n_2 + \cdots + n_r = n}$



each person gets items.

* ordering n items:

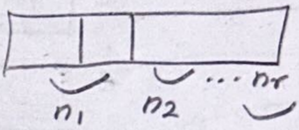


$n!$ ways

(or)

* Distribute them
* Ask them to order

  (n_1 arranges in n_1 slot)



* Distribute: subsets. [Combination]
* person $1 \to n_1!$ choices to order in n_1 slot

  $2 \to n_2!$

$$\therefore C \; n_1! \, n_2! \, n_3! \ldots n_r! = n!$$

$$\text{number of partition} = \frac{n!}{n_1! \, n_2! \, n_3! \ldots n_r!}$$

[multinomial co.eff]

Generalizes binomial case $\to r = 2, \; n_1 = K, \; n_2 = n - K$

But multinomial this is not the case $\to r = 3, \; n_1 = K_1, \; n_2 = K_2, \; n_3 = K_3$

---

52 (playing) card deck, fairly dealt to four players. P(each one gets an ace)

Solu:

outcomes:

$$\text{no. of outcomes: } \frac{52!}{13! \, 13! \, 13! \, 13!} = ?$$

52 cards

$\downarrow$

partition

$\downarrow$

13 cards each

(equally likely)

- Constructing an outcome with one ace for each person.

→ Distribute the aces (4 aces) → 4 ways, 3 ways, 2 ways, 1 way. (4!)

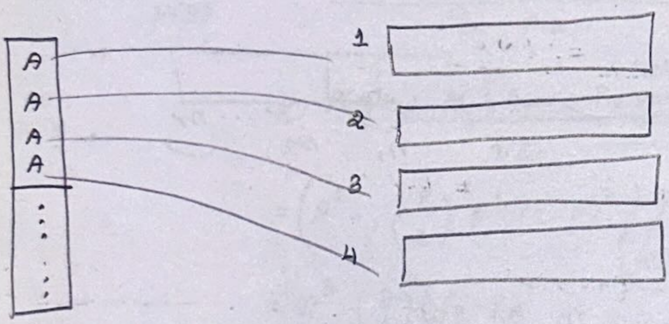
→ Distribute the remaining 48 cards.

$$\frac{48!}{12!\,12!\,12!\,12!}$$

$$\frac{4\times3\times2\times1 \cdot \dfrac{48!}{12!\,12!\,12!\,12!}}{\dfrac{52!}{13!\,13!\,13!\,13!}}$$

A smart solution

Stack the deck on top → Aces



Initially: $\frac{52}{52}$ slots to fill, $\frac{39}{52}$ (1 player got filled),

$\frac{26}{52}$ (2 players got filled)

$\frac{13}{52}$ (3 players got filled)

$= 0.105$

Assume: Already arranged (4 aces) → 1 → 4 choices $\frac{52}{52}$

2 → 3 choices $\frac{39}{51}$ → 1 already taken (cards)

Multinomial Probabilities

Generalization of binomial probabilities.

* Balls of different colors $i = 1, \ldots, r$
* Prob of picking a ball of color i is P_i.
* drawn n balls, independently.
* Given non-negative numbers n_i, with $n_1 + \cdots + n_r = n$
* Find: $P(n_1$ balls of color 1, n_2 balls of color $2, \ldots, n_r$ balls of color $r)$

Special case $r=2$; colors: 'heads', 'tails'.

n_1 of first color $\longrightarrow$ k heads $\qquad P_1 = P$

n_2 of second colour $\longrightarrow (n-k)$ tails $\qquad P_2 = 1-P$

$$P(k \text{ heads}) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

$n_1 = k$

$n_2 = n-k$

$$= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$\boxed{P(k \text{ heads}) = \frac{n_1!}{n_1!\, n_2!} P_1^{n_1} P_2^{n_2}}$$

$n=3$, $m=7$ balls

1 1 3 1 2 2 1

$\hookrightarrow$ 2nd ball has 1 color.

$4 \rightarrow$ 1st colour

$2 \rightarrow$ 2nd color

$1 \rightarrow$ 3rd color

type: $(4, 2, 1)$

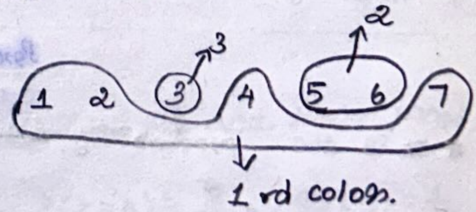
$\boxed{n_1 + n_2 + n_3 + \cdots = n}$

$$P(4, 2, 1) = P_1 P_1 P_3 P_1 P_2 P_2 P_1$$

$$= P_1^4 P_2^2 P_3$$

$$\boxed{P(\text{particular sequence of type } (n_1, n_2, \ldots, n_r)) = P_1^{n_1} P_2^{n_2} \cdots P_r^{n_r}}$$

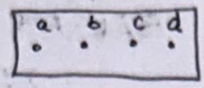
Sequence of type $(n_1, n_2, \ldots n_r)$ $\longleftrightarrow$ Partition of $\{1, \ldots, n\}$ in to subsets of sizes $n_1, n_2, \ldots, n_r$.



$$P(\text{get type } (n_1, n_2, \ldots, n_r)) = \frac{n!}{n_1!\, n_2! \cdots n_r!} P_1^{n_1} P_2^{n_2} \cdots P_r^{n_r}$$

$\longrightarrow$ No. of partitions $\times$ probabilities (for all sequence)

Random variable:

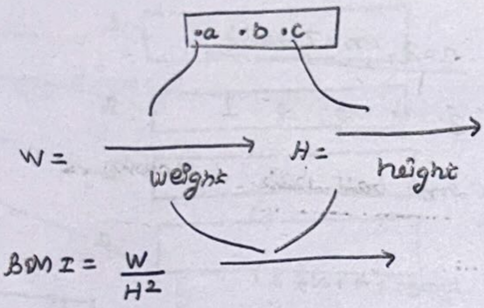


Select at random & measure the weight.

* weight is the function of the outcome of the experiment.
 (once outcome known – weight known)

Student $\longrightarrow$ $\boxed{W}$ $\longrightarrow$ weight

* A random variable (r.v) associates a value to every possible outcome.



'once o/p known W, H, BMI are known'

W, H, BMI are random variables.

Notation: X (r.v) – numerical value x

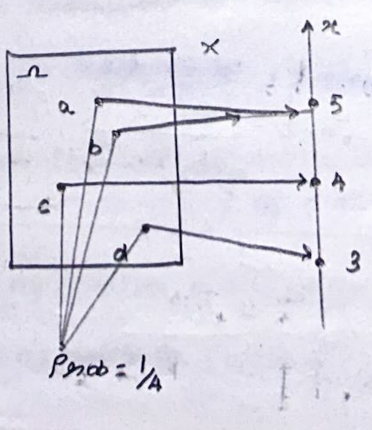
Random variables – Functions of sample space.

* several r.v can be assigned to a single Ω
* Combination of several r.v – also a r.v

– meaning of $X + Y$ = r.v takes $x+y$ when X takes x and Y takes y values.

Probability mass functions (X)

* It is the 'probability law' or 'distribution' of X.



$X = 5$, $\{w : X(w) = 5\} = \{a, b\}$

$$\boxed{p_X(x) = P(X = x)}$$

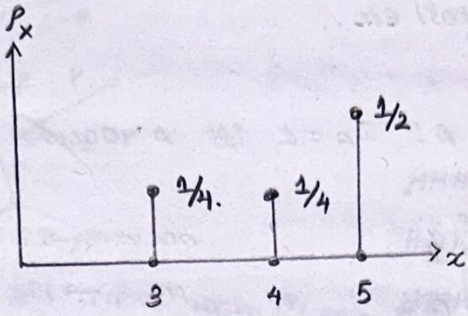
$P(X=5) = a+b$

$= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$= P(\{w \in \Omega$ s.t. $X(w) = x\})$

↓

outcome.



Properties: $p_X(x) \geq 0$

$$\boxed{\sum_x p_X(x) = 1}$$

PMF calculation (discrete)

two rolls of a tetrahedral die.

Y = second roll

Y \ X	1	2	3	4
4	5	6	7	8
3	4	5	6	7
2	3	4	5	6
1	2	3	4	5

X = first roll

→ 1/16 (likely) → each one's prob.

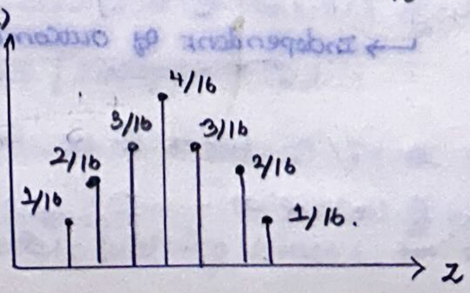
$Z = X + Y$

$p_Z(z) = ?$ (For all possible values)

* repeat for all z
* collect all possible outcomes for which $Z = z$ & add their prob.

$P_Z(Z=2) = 1/16$, $P(Z=3) = \frac{1}{8}$, $P(Z=4) = \frac{3}{16}$, $P(Z=5) = \frac{1}{4}$

$P(Z=6) = \frac{3}{16}$, $P(Z=7) = \frac{1}{8}$, $P(Z=8) = 1/16$



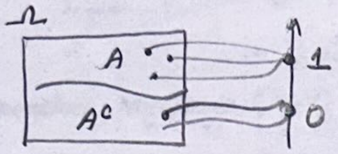
Simplest r.v : Bernoulli with parameter $p \in [0,1]$

$$X = \begin{cases} 1, & \text{w.p } p \\ 0, & \text{w.p } 1-p \end{cases}$$

* models a trial: results in success / failure

eg: Head or tail etc..

Indicator r.v of an event A: $I_A = 1$ iff A occurs

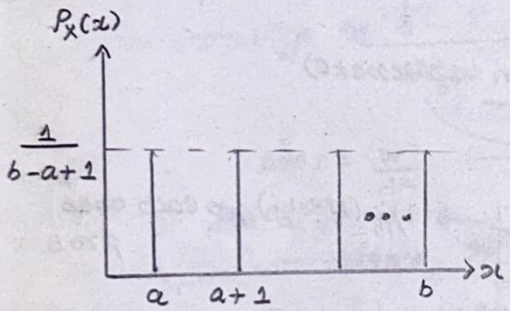


occurs → 1 (success)

No → 0 (failure)

$p_{I_{(A)}}(1) = P(I_A = 1) = P(A)$

Discrete uniform R.v (a, b : parameters)



(Equally likely)

Interval : $\{a, a+1, \dots b\}$

No. of elements = $b-a+1$ Possible values

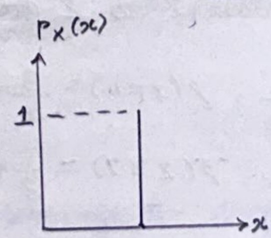
R.v $X : X(\omega) = \omega$

model: complete ignorance [Digital clock: 11:52:26 seconds]

$\{0, 1, \dots 59\}$

Each one is equally likely: $\frac{1}{60}$

Special case : a=b



only one value - No random in it.

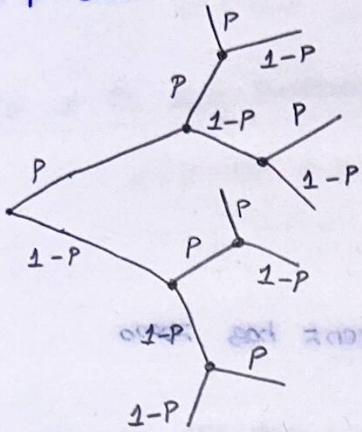
constant / deterministic r.v.

→ Independent of outcome [Only one value]

Binomial R.v

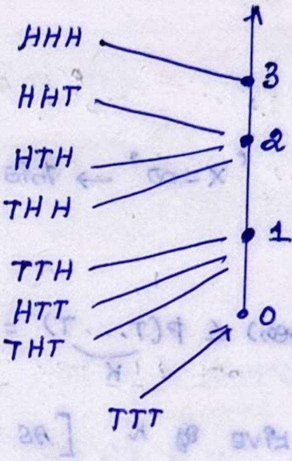
positive integers n; $P \in [0,1]$

* P - success, $P(\text{Heads}) = P$



HHH
⋮
TTT

length $= n = 3$



model: fixed numbers of identical & independent trial (no. of success)

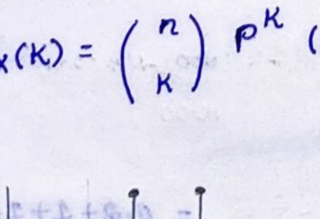
PMF

$$P_X(2) = P(X=2) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} \quad (\text{or}) = 3\left(P^2(1-P)\right)$$

$$= \binom{3}{2} P^2 (1-P)$$

Binomial formula.

$$P_X(K) = \binom{n}{K} P^K (1-P)^{n-K} \quad \text{for } K = 0, 1, \dots, n$$



$n = 3$

Geometric R.v

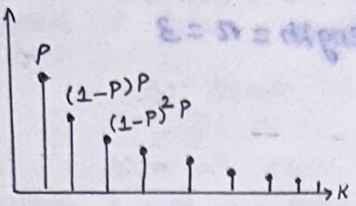
$P: 0 < P \leq 1$

* inly many tosses [Independently]
* $P(\text{Heads}) = \infty$

Model: waiting times; no. of trials until a success.

$P_X(k) = P(X=k) = P(\underbrace{T \ldots T}_{k-1}\overset{\downarrow}{H})$ k th trial

$= (1-p)^{k-1} p$ $\quad k = 1, 2, 3, 4, \ldots$



P(No head ever)

TTT..... 'X = ∞' → This particular event has zero probability.

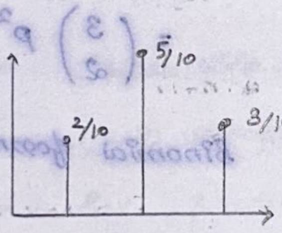
$P(\text{no heads ever}) \leq P(\underbrace{T \ldots T}_{k}) = (1-p)^k$

Irrespective of k [As $k \to \infty$ $(1-p) \to 0$]

Expectation / mean of R.V

motivation: game: 1000 times.
Random gain at each play

$$X = \begin{cases} 1, & \text{w.p } 2/10 \\ 2, & \text{w.p } 5/10 \\ 4, & \text{w.p } 3/10 \end{cases}$$



In 1000 times,

200 times → 1

500 times → 2

300 times → 4

Total gain (Average) $= \dfrac{1(200) + 2(500) + 4(300)}{1000}$

$= \frac{2}{10} + 1 + \frac{12}{10} = 0.2 + 1 + 1.2$
$= 2.4.$

$E[X] = \sum_x x P_X(x)$ ∴ we took each x and multiplied by its respective o/p prob.

Interpretation:
Average in large number of independent repetitions of experiment.

If we have an infinite sum, it needs to be well-defined.
we assume $\sum_x |x| P_X(x) < \infty$. [Converging]

'Sum - Finite'

Expectation of Bernoulli R.V:

$$X = \begin{cases} 1, & \text{w.p } p \\ 0, & \text{w.p } (1-p) \end{cases} \qquad E[X] = 1 \cdot p + 0 \cdot (1-p) = p$$

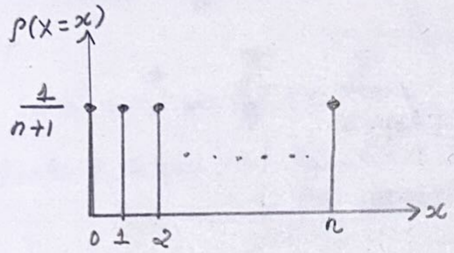
If X is the indicator of an event A, $X = I_A$:

$X = 1$ iff A occurs: $p = P(A)$

$$E[I_A] = P(A).$$

Expectation of a uniform R.V

★ Uniform on $0, 1, \ldots, n$



$\therefore \dfrac{1}{b-a+1} = \dfrac{1}{n-0+1} = \dfrac{1}{n+1}$

$$E[X] = \sum_x x\, P_X(x) = 0\left(\frac{1}{n+1}\right) + 1\left(\frac{1}{n+1}\right) + \cdots + n\left(\frac{1}{n+1}\right)$$

$$= \left(\frac{1}{n+1}\right) \cdot (1+2+\cdots n)$$

$$= \left(\frac{1}{n+1}\right) \frac{n(n+1)}{2}$$

$$= \frac{n}{2}$$ (finite) $\rightarrow$ midpoint of the x values

'symmetric'

Expectation of a population average.

★ n students: weight of ith student: x_i and pick a student at random, all equally likely.

Soln:

X: weight of the students (Assume x_i - distinct) $\rightarrow$ unique for a student.

$P_X(x_i) = \dfrac{1}{n}$ [choosing a student among n]

$$E[X] = \sum_i x_i \frac{1}{n} = \frac{1}{n} \sum x_i$$

Gives the true average over the entire population.

Elementary properties of expectation

$$E[X] = \sum_{x} \underbrace{x}_{\geq 0} \underbrace{P(X=x)}_{\geq 0}$$

★ If $X \geq 0$, then $E[X] \geq 0$ [Non negative]

★ If $a \leq X \leq b$, then $a \leq E[X] \leq b$

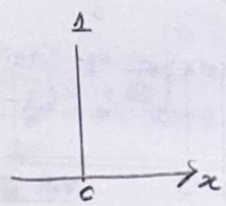
Proof:

$$E[X] = \sum_{x} x \, P_x(x)$$
$$= a \sum_{x} P(X=x)$$
$$= a(1)$$

$$E[X] = \sum_{x} x \, P(X=x)$$
$$= b \sum_{x} P(X=x)$$
$$= b(1)$$

If c is a Constant $E[c] = c$

$$E[X] = \sum_{x} x \; P(X=x)$$
$$= c \sum_{x} P(X=x) = c$$



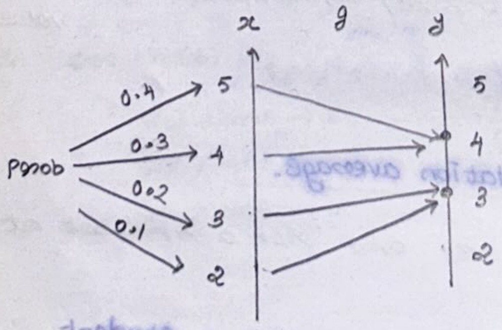
$(b = a = c \text{ (constant)})$

only one term $P(c) = 1$

one point $X = c$

The expected value rule, for calculating $E[g(x)]$

Let X be a R.V and let $Y = g(x)$



Averaging over y: $E[Y] = \sum_{y} y \, P(Y=y)$

$$= 3\,P(Y=3) + 4\,P(Y=4)$$
$$= 3(0.1+0.2) + 4(0.3+0.4)$$
$$= 3(0.3) + 4(0.7)$$
$$= 3.7$$

Averaging over x:

$$3(0.1) + 3(0.2) + 4(0.3) + 4(0.4) = 3.7$$

From here.

when 10% of $x = 2$, $y = 3$

20% of $x = 3$, $y = 3$

$$E[Y] = E[g(x)] = \sum g(x) P_X(x)$$

Proof: $E[g(x)] = \sum g(x)\, P_X(x)$

going through all

$$= \sum_{Y} \sum_{x : g(x) = Y} g(x)\, P_X(x) \quad \text{[Fixing a } Y\text{]}$$

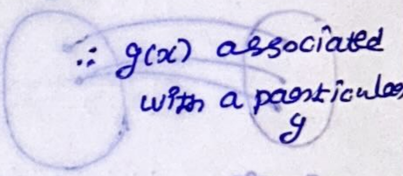
From diag:

For any Y — consider all x then go through all Y.

$$= \sum_{Y} \sum_{x : g(x) = Y} g(x)\, P_X(x)$$

(x associated with Y)

$$= \sum_{Y} \sum_{x : g(x) = y} y\, P_X(x)$$



$$= \sum_{Y} y \sum_{x : g(x) = Y} P_X(x)$$

$$= \sum_{Y} Y \sum_{x : g(x) = Y} P_X(x)$$

↳ Fixing y then go through x gives prob of that y

$$= \sum_{Y} Y\, P_Y(y)$$

$$= \sum_{Y} y\, P_y(y)$$

$$= E[Y]$$

Note: $E[x^2] = \sum_{x} x^2 P_X(x)$

$g(x) = x^2$

$\ast$ $E[x^2] \neq [E[x]]^2$ → Precaution.

$\therefore E[g(x)] \neq g[E[x]]$ → we can't able to change order.

Linearity of expectation

$E[ax+b] = aE[X]+b$

$X \to$ salary   case: gets double the salary + bonus.

$E[X]$ = average

$g(x) = ax+b$

$E[X] = \sum_x x\, P_X(x)$

$E[Y] = \sum_x g(x)\, P_X(x)$

$E[ax+b] = \sum_x (ax+b)\, P_X(x)$

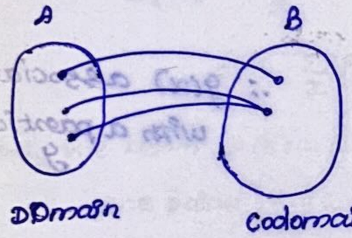
$= a\sum_x x P_X(x) + b\sum_x P_X(x)$

$= aE[X] + b(1)$

$= aE[X]+b$

$E[g(x)] \neq g(E[X]) \longrightarrow$ Exceptional

Functions



* Every element of A is mapped to B.
(may not be unique)

$f: A \longrightarrow B$
(maps elements of A to B)

$x \in A,\ y \in B$
each $x \in A$, appears in exactly one pair.

$A = \{-1, 0, 1, 2\}$

B = Real numbers

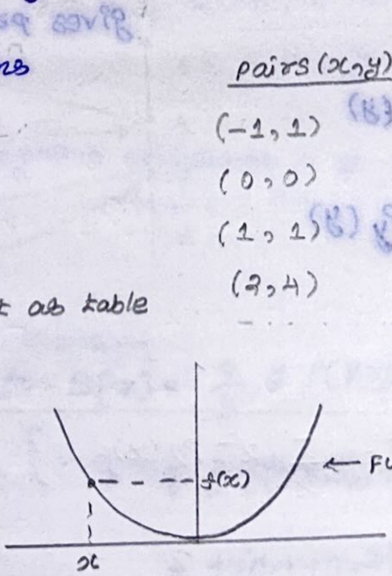
f: 'squares'

pairs (x, y)

$(-1, 1)$
$(0, 0)$
$(1, 1)$
$(2, 4)$
...

$\{(x, x^2) \mid x \in \mathbb{R}\}$
$\in \mathbb{R}^2$
Real numbers

* we can represent it as table



← Function : entire plot.

Notation:

$f: R \to R$ defined by $f(x) = x^2 \longrightarrow$ Formal

function $x^2 \longrightarrow$ Informal.

$f: R \to R$ defined by $f(z) = z^2$ $\rightarrow$ function z^2

General:

the function f $\rightarrow$ Formal

the function $f(x)$ $\rightarrow$ Informal

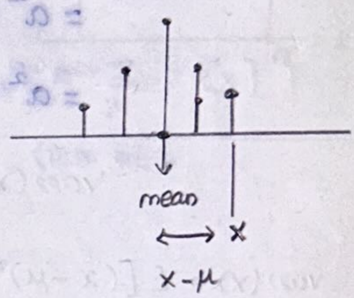
$$\xrightarrow{x} \boxed{f} \longrightarrow f(x)$$

Lecture-6 – Discrete R.V – II

Variation: Amount of randomness – distribution of pmf.
(Spread of PMF)

* Random variable X, with mean $\mu = E[X]$

* Average distance from the mean?



$$E[X-\mu] = E[X] - \mu$$
$$= \mu - \mu$$
$$= 0 \rightarrow \text{Informative}$$

Average absolute value of the distance from the mean:

$$\text{var}(X) = E\left[(X-\mu)^2\right] \rightarrow \geq 0 \rightarrow \text{non negative.}$$

$$E[g(X)] = \sum_x g(x)\, P_X(x)$$

$$\text{var}(X) = E[g(x)]$$
$$= E\left[(x-\mu)^2\right]$$
$$= \sum_x (x-\mu)^2\; P_X(x)$$

$\therefore$ units of mean & variance are different.
$m \neq m^2$

Standard deviation: $\sigma_X = \sqrt{\text{var}(X)}$

Properties

$\mu = E[X]$

$Y = X + b$

$E[Y] = E[X] + b$
$= \mu + b$

$\boxed{\therefore E[Y] = \mu + b}$

$\text{var}(aX+b) = a^2\, \text{var}(X)$

$\text{var}(Y) = E[(Y - \gamma)]$
$\downarrow$
Y's mean

$= E\left[(X+b-\gamma)^2\right]$

$= E\left[(X+b-(\mu+b)^2\right]$

$= E[(x-\mu)^2]$

$= var(x) \longrightarrow$ where $a = 1$

Let $Y = aX$:

$\gamma = E[Y] = a\, E[x] = a\mu$

$var(y) = E[(ax - a\mu)^2]$

$= E[a^2(x-\mu)^2]$

$= a^2 E[(x-\mu)^2]$

$= a^2 var(x)$

$$var(x) = E[x^2] - (E[x])^2$$

$var(x) = E[(x-\mu)^2]$

$= E[(x^2 - 2x\mu + \mu^2)]$

$= E[x^2] - 2E[x\mu] + E[\mu^2]$

$= E[x^2] - 2\mu E[x] + \mu^2$

$= E[x^2] - 2(E[x])^2 + (E[x])^2$ $\quad \therefore \mu = E[x]$

$= E[x^2] - [E[x]]^2$

Variance of the bernoulli & uniform

$x = \begin{cases} 1, & \text{w.p. } P \\ 0, & \text{w.p. } 1-P \end{cases} \longrightarrow$ Bernoulli $\quad \therefore E[x] = p$ (Bernoulli)

$var(x) = \sum_x (x - E[x])^2\, P(x = x)$

$= (1-P)^2 P + (0-P)^2(1-P)$

$= P - 2P^2 + \cancel{P^3} + P^2 - \cancel{P^3}$

$P - P^2$

$= P(1-P)$

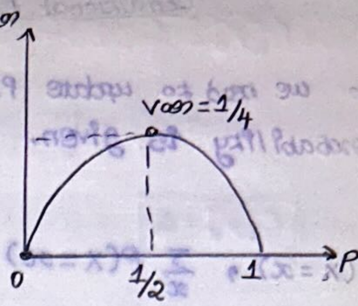
$var(x) = E[x^2] - (E[x])^2$

$\therefore$ when $x = 0$, $x^2 = 0$
$x = 1$, $x^2 = 1$

$= E[x] - E[x]^2$

$= P - P^2$

$= P(1-P)$



$var(P=0) = var(P=1) = 0$

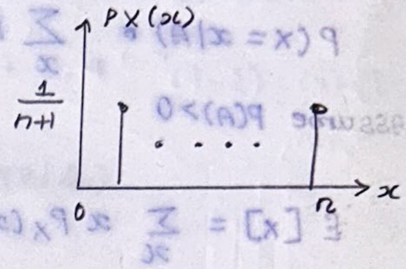
↳ Random

As var $= 1/4 \rightarrow$ when $P = 1/2$

we can't say it's fair.

Variance → Biggest if it is fair.

Variance of uniform



$$var(X) = E[X^2] - (E[X])^2$$

$$\sum_x x P_x(x) = \left(\frac{1}{n+1}\right) 0 + \left(\frac{1}{n+1}\right) 1^2 + \cdots$$

$$= \left(\frac{1}{n+1}\right)(0^2 + 1^2 + 2^2 + \cdots + n^2)$$

$$var(X) = \left(\frac{1}{n+1}\right)(0^2 + 1^2 + \cdots + n^2) - \left(\frac{n}{2}\right)^2$$

$$= \left(\frac{1}{n+1}\right)\left(\frac{1}{6} n(n+1)(2n+1)\right) - \left(\frac{n}{2}\right)^2$$

$$= \frac{1}{6} n(2n+1) - \frac{n^2}{4}$$

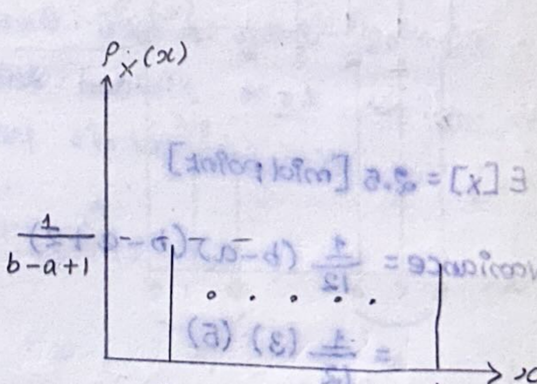
$$= \frac{2n^2}{6} + \frac{1}{6} n - \frac{n^2}{4}$$

$$= \frac{4n^2 - 3n^2}{12} + \frac{2}{12} n$$

$$= \frac{n^2}{12} + \frac{2}{12} n$$

$$= \frac{1}{12} n(n+2)$$

suppose ←



'shifting a PMF' – adding a constant variable.

$\boxed{b = n + a}$ $\quad n = b - a$

Adding a constant doesn't changes the variance

$$var(X) = \frac{1}{12}(b-a)(b-a+2)$$

Conditional PMFs & Expectations on an event

So we need to update PMF to a conditional PMF when a conditional probability is given.

$P(X=x)$, $\sum_x P(X=x) = 1$

Condition on an Event $A \rightarrow$ use conditional probabilities

$P(X=x|A)$ • $\sum_x P_{X|A}(x) = 1$

Assume $P(A)>0$

$E[X] = \sum_x x P_X(x)$, $E[X|A] = \sum_x x P(X=x|A)$

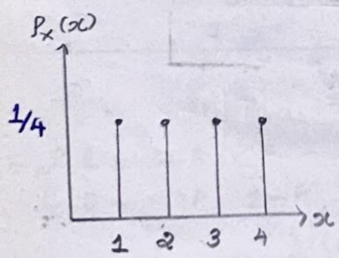
(normal)

Conditional probability model is another probability model applying to new situation.

$E[g(x)] = \sum_x g(x) P(X=x)$

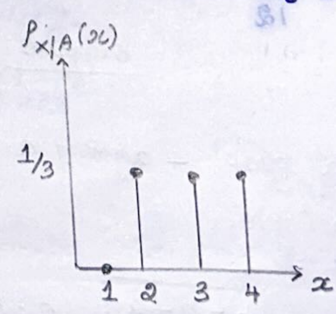
$E[g(x)|A] = \sum_x g(x) P_{X|A}(x)$

Example of conditioning



Let $A = \{X \geq 2\}$

$E[X] = 2.5$ [mid point] $\longrightarrow$



Variance $= \frac{1}{12}(b-a)(b-a+2)$

$= \frac{1}{12}(3)(5)$

$= \frac{5}{4}$

$E[X|A] = 3$

$\text{Var}[X|A] = \frac{1}{12}(2)(4)$

$= \frac{8}{12} = \frac{2}{3}$

$g(x) = (x-\mu)^2$

$var[X|A] = \sum_x g(x) P_X[x]$

$= \frac{1}{3}(4-3)^2 + \frac{1}{3}(3-3)^2 + \frac{1}{3}(2-3)^2$

$= \frac{2}{3}$

"In conditional model — Less randomness"

Total expectation theorem

→ "Allow as to divide & conquer"

Total prob theory: $P(B) = P(A_1)P(B|A_1) + \dots + P(A_n)P(B|A_n)$

Let,

$B = \{X = x\}$

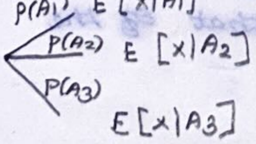
$P(X=x) = P(A_1) P_{X|A_1}(x) + P \dots + P(A_n) P_{X|A_n}(x)$

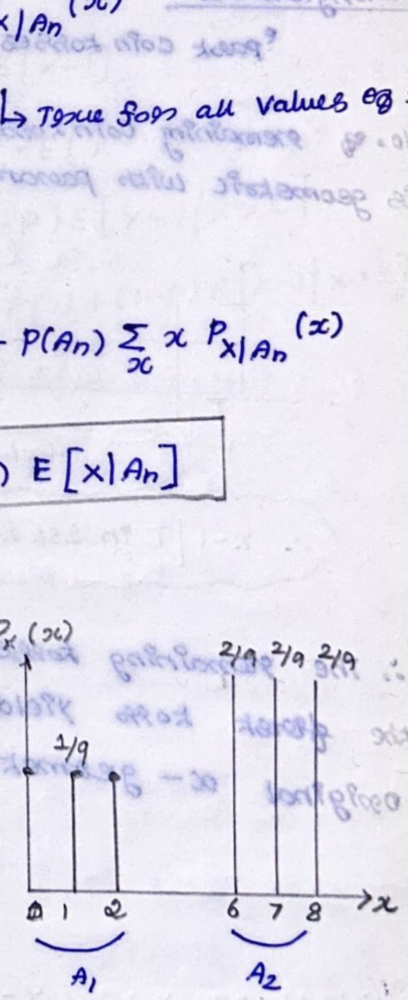
↳ True for all values of x.

mul by x:

$$\sum_x x P_X(x) = P(A_1) \sum_x x P_{X|A_1}(x) + \dots + P(A_n) \sum_x x P_{X|A_n}(x)$$

$$\boxed{E[X] = P(A_1)E[X|A_1] + \dots + P(A_n)E[X|A_n]}$$





$P(A_1) = 1/3$, $P(A_2) = 2/3$

$E[X|A_1] = 1$ [∴ A_1 happened]

X is certainly from A_1 [midpoint = 1]

$E[X|A_2] = 7$ [midpoint]

$$E[X] = \frac{1}{3}\cdot 1 + \frac{2}{3}\cdot 7$$

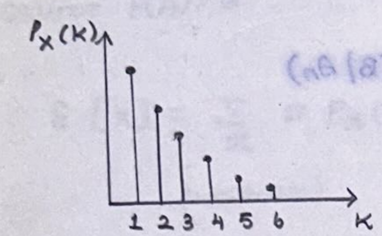
$$= \frac{1}{3} + \frac{14}{3} = \frac{15}{3} = 5$$

$$\boxed{E[X] = 5}$$

Geometric PMF : Memorylessness & Expectation

X: number of independent coin tosses until first head; $P(H) = p$

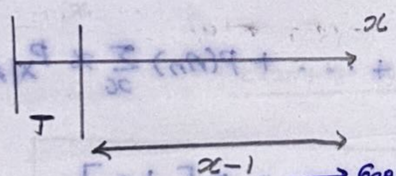
$$P_X(k) = (1-p)^{k-1}\, p, \quad k = 1, 2, \ldots$$



Memorylessness:

'past coin tosses doesn't affect the future'

No. of remaining coin tosses, conditioned on tails in the first toss is geometric with parameter, p.



→ Geometric with (P)

∴ X−1 | T in 1st toss

∴ The remaining tosses Geometric probability in that the first toss yielded tail is same as that of the original x− geometric probability.

'Memorylessness'

Conditioned on $X > 1$, $X - 1$ is geometric with parameter p

eg

$$P(X-1 = 3 \mid x > 1) = P(T_2 T_3 H_4 \mid T_1)$$

'Independent' — conditional = unconditional

$$= P(T_2 T_3 H_4)$$

$$\boxed{P_{X-1|X>1}(3) = (1-P)^2 P = P_X(3)}$$

Generalize

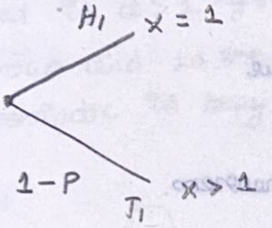
$$P_{X-1|X>1}(k) = P_X(k)$$

Conditioned on $X > n$, $X - n$ is geometric with P

$$P_{X-n|X>n}(k) = P_X(k) = P_{X-n|X>n}(k)$$

$$E[X] = \sum_{k=1}^{\infty} k\, P_X(k) = \sum_{k=1}^{\infty} k(1-P)^{k-1} P$$

Simplify:



[Total prob theorem]

$$E[X] = 1 + E[X-1]$$

$$= 1 + P \cdot E[X-1 \mid X=1] + (1-P)\, E[X-1 \mid X>1]$$

$$= 1 + P(0) + (1-P)\, E[X-1 \mid X>1]$$

Lecture: 17

Markov chains II

"In real world anything can be appr modelled by markov chains"

Discrete state, Discrete time, time-homogeneous:

* Transition prob P_{ij}

$\downarrow$ $\hookrightarrow$ next state

current

model: Time invariant

$$i \xrightarrow{P_{ij}} j$$

Same everytime (P_{ij}) $\rightarrow$ time-homogeneous

$$r_{ij}(n) = P(X_n = j \mid X_0 = i)$$

Key recursion:

$$r_{ij}(n) = \sum_k r_{ik}(n-1) P_{kj} \rightarrow \text{Always true}$$

$\downarrow$

Divide & conquer: Total prob theorem.

fix i $\rightarrow$ different ways to get j

$\downarrow$

$\boxed{\text{Different (states } k)}$ $\rightarrow$ prob you see yourself at previous state

P_{kj} $\rightarrow$ prob k to j.

(prev diagram): Warm up

Structural properties:

$$P(X_1 = 2, X_2 = 6, X_3 = 7 \mid X_0 = 1) = P_{12} \cdot P_{26} \cdot P_{67} \quad \text{(Trajectory)}$$

$2 \rightarrow 6 \rightarrow 7$

$P(X_4 = 7 \mid X_0 = 2) =$ $\rightarrow$ use recursion

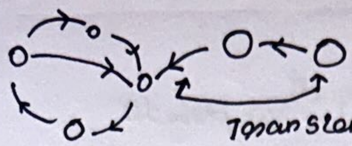
(or)

(4 steps X_7)

$$= P_{26} P_{67} P_{76} P_{67} + P_{26} P_{66} P_{66} P_{67} + P_{21} P_{12} P_{26} P_{67}$$

As index goes bigger, the job becomes huge.

'Recursion - Clever way - Not exponential time horizon'



Recurrence

* Entered into a trap after leaving from transient state.
* Two recursion (without connection) → previous eg.
↓
Initial condition makes the difference.

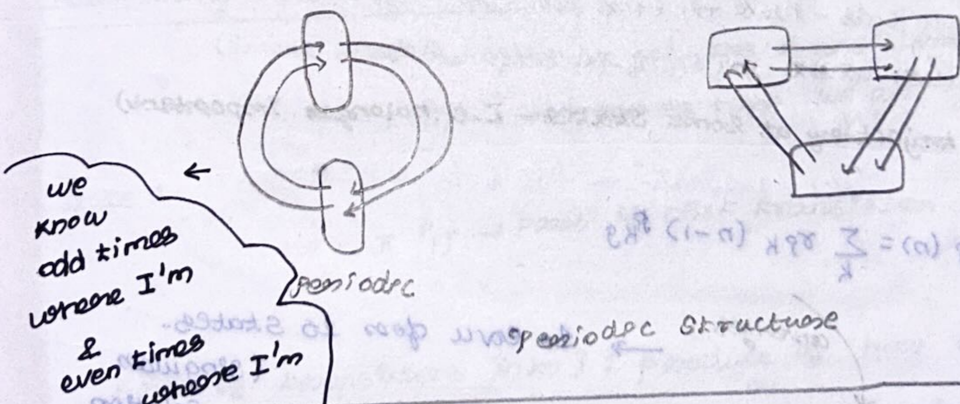
---

* with a single recurrent classes - Inital cond doesn't matter
* with more recurrent classes — I.C matters

<u>Another way I.C matters</u>

* periodic nature

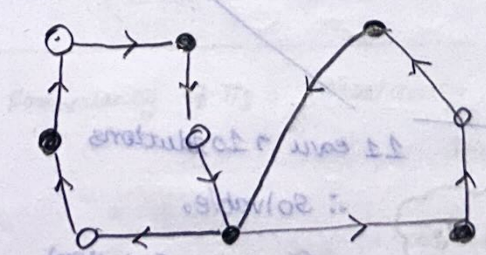
Grouped if $d > 1$ groups (→ clusters) so that all transitions from one group lead to the nex group. (States in a recurrent class are periodic if they can be grouped like → above



---

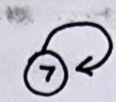
periodicity: osciluate

<u>Periodic or not?</u>



periodic - From colored to non-coloured.

Not periodic: If you have a self transition



'In self transition: you may be in your own reign' — Not periodic

---

Steady - state probabilities

Do $r_{ij}(n)$ converge to some x_j? (Independent of the initial state.)

Solu:

Yes if,

→ Recurrent states are all in a single class (Initial State doesn't matters)

→ single recurrent class is not periodic.

↑ Steady - state Markov theorem

Assuming yes (converge):

a recurrence markov chain

↗

From two ends - move → some point we may meet at same point,

Collide → No matter where you start. (Collided - Agreed start).

(Same trajectory at some states - I.C no longer important)

$$r_{ij}(n) = \sum_k r_{ik}(n-1)\, P_{kj}$$

As $n \to \infty$ — converges → 10 eqn for 10 states. (Singular solution eqn) ↓ Doesn't have unique solution

$$\pi_j = \sum_k \pi_k P_{kj},$$

for all j

$\pi_k \to$ steady state probability

Additional eqn:

$$\sum_j \pi_j = 1$$

↓ prob of different states

$\therefore r_{ij}(n)$ converge to some π_j

11 eqn, 10 solutions

∴ Solvable.

(Unique Solution)

↓

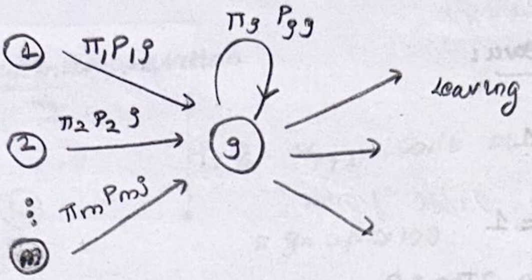
... we can find steady state solution prob of markov.

Visit frequency interpretation

Coin: how frequency can I get a head.

$\pi_j = \sum_k \pi_k P_{kj}$ → Balance equations.

* (Long run) freq of being in j : π_j → where i can find j (over a long run)
* Frequency of transitions $k \to j$: $\pi_k P_{kj}$
* Freq of transitions into j : $\sum_k \pi_k P_{kj}$



particular state (j) : Transition into j (vs) Transition starting from j (Includes self transition)

Interpreting π_j as frequency:

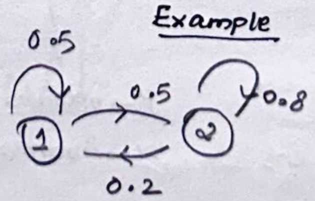
* How often we get the transitions into j
* π_1 → Fraction of time we are going to be at state 1.
* P_{1j} → prob of next transition into j.

Freq of transitions into j : product of how often the π_p will be in that state × transition of that in to j.

$$\boxed{\pi_j = \sum_k \pi_k P_{kj}}$$ → sum of all $\pi_1 P_1 j$, $\pi_2 P_{2j}$, $\pi_m P_{mj}$

Frequency of π_j = Frequency with which transition into j happens.

Example



$\pi_1 = \pi_1(0.5) + \pi_2(0.2)$
↓
I was at State 1,
then made self transition

$\pi_2 = \pi_1(0.5) + \pi_2(0.8)$
↓ Transition ↓ self transition.

$\boxed{0.5\pi_1 = 0.2\pi_2} \rightarrow 1$ $\boxed{0.5\pi_1 = 0.2\pi_2} \rightarrow 2$

Both eqn telling the same

Linearly dependent eqn:

$$\sum_K \pi_K = 1$$

$$\pi_1 + \pi_2 = 1$$

$$0.5\pi_1 - 0.2\pi_2 = 0$$

$$\pi_1 = \frac{2}{7}, \quad \pi_2 = \frac{5}{7}$$

'keep visiting both states'

Stable state prob of two diff probabilities.

Prob: of finding at state 1 = 2/7, at state 2 = 5/7

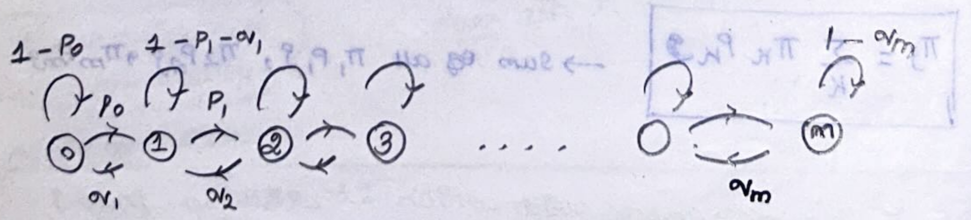
'No matter where you stand'

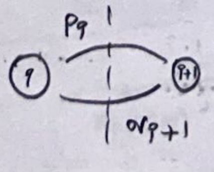
'Not-periodic' — 'self recurrence'.

Special example

model: lead you to better understanding

Birth-death process





$\pi_q P_q = \pi_{q+1} \alpha_{q+1}$

P → upward transition
α → downward transition

Solu: * ↑ or ↓ or nothing happens

Assuming constant rates

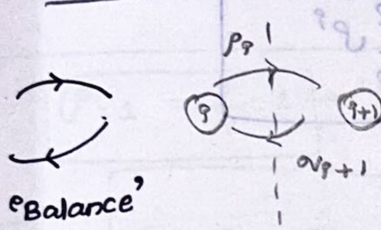
$p_i = p$, $v_i = v$ (As in supermarket.

$\rho = p/v$ = load factor

<u>Good model for Server Systems</u>

* Phone line
* Hospital beds.
* Flu spread (more people: high prob of spread)
(New incidence of disease)

<u>Shortcut</u>: Frequency interpretation



p_i & v_{i+1} can't be much different

* go up ≈ 100
* go down ≈ 99

Balance: No. of upward transitions can't be much different from downward transitions.

'Birth ≈ Death rate'

$$\boxed{\pi_i p_i = \pi_{i+1} \cdot v_{i+1}}$$

↓ Recursion ↓

If π_i is known π_{i+1} also.

∴ Two flow are same

$$\boxed{\therefore p_i = p, \ v_i = v}$$

$\boxed{\text{Unknown: } \pi_0}$

$\sum \pi_i = 1$ → we can find all of the π_i s.

$\boxed{\text{'}\pi_i p_i = \pi_{i+1} \cdot v_{i+1}\text{'} \rightarrow \text{Express in } \pi_0}$

Then solve.

$$\pi_{i+1} = \pi_i \left(\frac{p}{v}\right)$$

$\frac{p}{v} = \frac{\text{going up}}{\text{going down}}$ = How loaded

$$\boxed{\pi_{i+1} = \pi_i \rho}$$

$\boxed{p = v}$ → balanced
moves likely in both direction.

when P is larger: The servers may get loaded (queues buildup)

$\circ \rightarrow \circ \rightarrow$ This direction as P is large

Drain down

$\pi_{i+1} = \rho \pi_i$

$\pi_1 = \pi_0 \rho$

$\pi_2 = \pi_1 \rho = \pi_0 \rho^2$

$\pi_3 = \pi_2 \rho = \pi_0 \rho^3$

$\boxed{\pi_i = \pi_0 \rho^i}$ → gsho

$$\sum_{i=0}^{n} \pi_i = 1$$

$$\sum_{i=0}^{n} \pi_0 \rho^i = 1$$

$$\boxed{\pi_0 = \frac{1}{\sum_{i=0}^{m} \rho^i}}$$

So we found steady state prob of all the states

If $\rho = 1$

$$\boxed{\pi_0 = \frac{1}{\sum_{i=0}^{m} 1} = \frac{1}{m+1}}$$

$\boxed{\pi_i = \pi_0}$

All s.s probability are equal: equally likely in long run.

"Symmetric random walk - walk of drunken man"

No bias (He may go left or right)

Equally likely in all states (He maybe anywhere)

If m is a huge number.

Supermarket: large place. $(\rho < 1)$

Customers served faster than they arrive

$$\pi_0 = \frac{1}{\left(\sum_{i=0}^{m} \rho^i\right)_{m\to\infty}} = \frac{1}{\sum_{i=0}^{\infty} \rho^i} = \frac{1}{\left(\frac{1}{1-\rho}\right)} = 1-\rho$$

Geometric Series

$$\boxed{\pi_i = (1-\rho)\rho^i}$$