Half course

To model some uncertainity.

Probability & Statistics

Girish Varma.

Lecture 1

- 12 lectures total, 2 lectures per week.
- Problem solving is required & have to be regular in class & also refer to reading material
- Weekly Quiz. (from Basic & Reading Material) on Moodle.
- 2 longer Quiz (of 1 hour each)

- In ML we use probability for prediction rather than modelling some uncertainity.

it can be superset of all possible outcome also

Sample space: Set of all possible outcomes

$\{H,T\}$ or $\{0,1\}$ in case of Toss coin.

Throw of dice: $\{1,2,3,4,5,6\}$ or $\{0,1,\ldots 5\}$ or $\{3,4,\ldots 9\}$

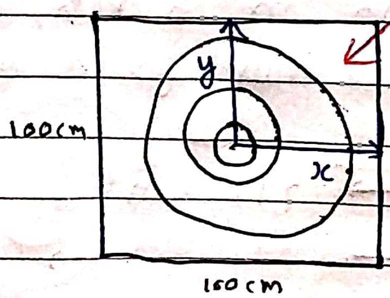
Discrete sample space as we know all possible outcome

- Keep on tossing until "head" comes

here sample space will be $\{0,1\} \times \{0,1\} \times \{0,1\} - - - - - \{0,1\}$

here sample space is superset of all possible outcome

- For Dart Board



we are allowed to throw dart outside ~~circle but inside circle~~.

here sample space will be

$\{(x,y): -50 \le x, y \le 50\}$

here sample space is superset of all possible outcome.

Sample space can also be written as

$A = \{-50 \le x \le 50\}$

$S = A \times A = A^2$

- Toss a coin 5 times : $\{0,1\}^5$

- In dart problem if we are not allowed to throw outside dart board circle.

$$\Omega = \{(r,\theta) : 0 \le r \le 50,\ 0 \le \theta \le 2\pi\}$$

- Shuffling a deck of cards.

Sample space for 1 card: $\{1, 2, \ldots 10, A, J, K, Q\} \times \{\diamondsuit, \heartsuit, \clubsuit, \spadesuit\}$

- But sample space for shuffled deck of card.

$$\sigma : \underbrace{\{1, 2, \ldots 52\}}_{\text{cards \#}} \xrightarrow{\text{mapped to}} \underbrace{\{1, 2, \ldots 52\}}_{\text{position \#}}$$

b/wn card# & position

so here sample space is a one-one & onto function

$$\Omega = \{\sigma : \{1, \ldots 52\} \to \{1 \ldots 52\} : \sigma \text{ is } \underline{\text{one-one}} \text{ fun}^n\}$$

$\therefore |\Omega| = 52!$ ← # of one-one function

Since size of both set are same & they are onto there for it will be onto also

Note: one-one, onto function, bijection, permutation all are same.

- Balls & Bins

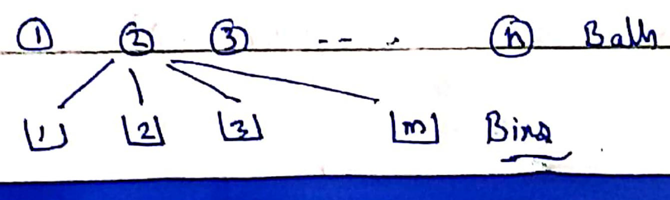
n balls are thrown in m buckets, buckets can be empty also, every ball must go in one bin.

$$\Omega = \{\overset{\text{function}}{f} : \{1, \ldots n\} \to \{1 \ldots m\}\} \quad \text{or}$$

$$\times\times\ \Omega = \{f : \{1 \ldots m\} \to \{1 \ldots n\}\}$$

↳ Here every Bin must have one Ball but that is not the case, as every element in domain must have image in a function

$$|\Omega| = m^n$$



- Random Undirected graph on n-vertices,

$$\# \text{edges} = \binom{n}{c} = \Big| \underbrace{\{ (a,b) : a \neq b \ \& \ a,b \in \{1, \ldots n\} \}}_{K} \Big|$$

any subset of 2.

$$\Omega = \{ S \subseteq K \}$$

$$\therefore \ |\Omega| = 2^{\binom{n}{2}}$$

i.e $\Omega = \{0,1\}^{\binom{n}{2}}$

Don't take. Take!

{Characteristics vector of $S \subseteq K$}

if $|K| = p$
then $\Omega = \{0,1\}^{p}$

subset of sample space.

Probabity of an event

- Toss 2 coin, both occurance are same.

Event $E \subseteq \Omega$, $E = \{00, 11\}$ every element in sample space has same probability

sample is finite or countabley infinite

It is a discrete & uniform probability then

Prob. of E, $|P(E)| = |E| / |\Omega| = 2/4 = 1/2$

$$P(E) = \frac{|E|}{|\Omega|} = \frac{2}{4} = 1/2$$

$$P(\Omega) = 1, \quad P(\phi) = 0$$

$$P(E) = \sum_{x \in E} P(x) = \frac{|E|}{|\Omega|}$$

works with finite only

this cannot be used with continious things (uncountable ∞ things)

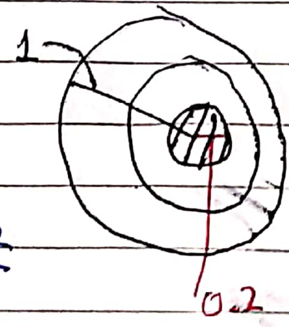
works for Countable ∞ & finite also

Biased coin $(2/3, 1/3)$

$$P(E) = \sum_{x \in E} P(x) = \frac{2}{3} \times \frac{2}{3} + \frac{1}{3} \times \frac{1}{3} = 5/9$$

Dart Board

In continious space we use area to find prob.



E : hit shaded portion

$$P(E) = \frac{\text{Area of shaded}}{\text{Area of whole}} = \frac{\pi \times 0.2 \times 0.2}{\pi}$$

$$= 0.04$$

• In case of 3-D space we will use volume to find prob. & to find area we will need to do integration

$P((0.1, 0.2)) = 0$, in continious space
cannot be a true at all

$P((0.1, 0.5), (1,2), (0.02, 0.01), (5, 1)) \stackrel{?}{=} 0$

$P(\Omega)$

$\sum_{x \in \Omega} p(x,y) = \sum_{x \in \Omega} 0 = 0$ wrong since, $P(\Omega) = 1$

∴ we cannot use "Σ" on continious Space, we have to use Integration.
"Σ" can be used in discrete space only.

• Toss a coin 5 times, what is the probability that you see 101 consecutively.

– $\Omega = \{0,1\}^5$, $|\Omega| = 32$
$E = \{ x \in \{0,1\}^5 : x$ has 101 as substring$\}$

$*101* : 4/32 \quad E_1$
$+ \ 101** : 4/32 \quad E_2$
$+ \ **101 : 4/32 \quad E_3$
$- \ 10101 : 1/32$

$P(E) = P(E_1 \cup E_2 \cup E_3)$
$= P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3)$
$- P(E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_3)$
$= 4/32 + 4/32 + 4/32 - 0 - 0 - 1/32 + 0$
$= 11/32$

Imp Approach

Divide complicated event in simpler event & calcute it's probabity & solve the complicate event. like previous example

TA Session Imp Notes only

- Difference Rule

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

- Boole's Inequality

$$P(A \cup B) \leq P(A) + P(B)$$

- Monotonicity.

$$P(A) \leq P(B), \quad \text{if} \quad A \subseteq B \text{ (if A is subset of B)}$$

- Uniform Prob. Space
  Unbiased coin toss, both H & T have same prob.
  but Biased coin does not has Uniform Prob. space

- Infinite probabity space.
  - Events where sample space is countable infinite
  so, Selecting a point in a circle is not an example of Infinite probability space.
  You cannot find all the possibilities in uncountably ∞
  $\therefore$ we avoid that case.
  So correct example of Infinite probability space will tossing a coin until simultaneous Head appears twice.

Four step Method

- 1 coin is tossed thrice, E = getting all the three face value as same i.e HHH or TTT

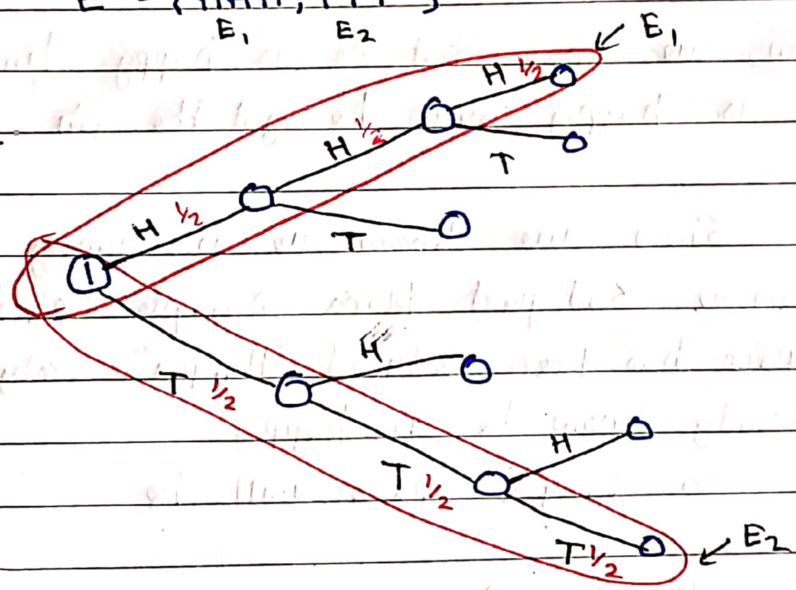
Step 1) Find the sample space

$|S| = 8$, $S = \{HHH, HHT, HTH, THH, -- TTT\}$

Step 2) Find Event of interest

$E = \{\underset{E_1}{HHH}, \underset{E_2}{TTT}\}$

Step 3)

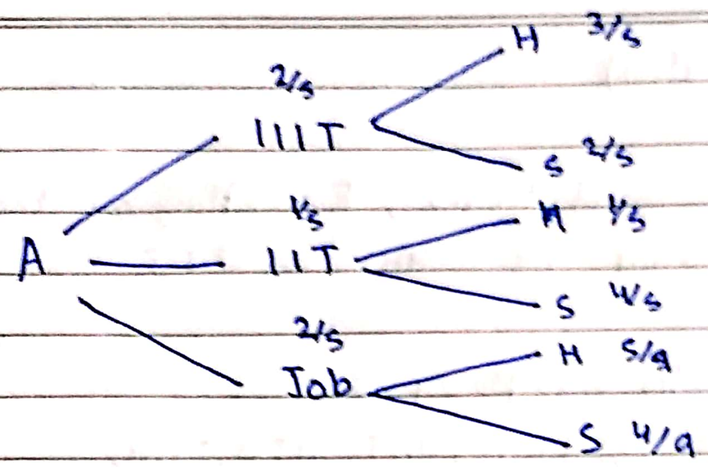


Step 4)

$P(E) = P(E_1) + P(E_2)$

$= \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$

- A person wants to join IIIT, IIT or Job with prob $2/5$, $1/5$, $2/5$ & prob that he will be happy or sad if he joins IIIT, IIT, or Job will be $(3/5, 2/5)$, $(1/5, 4/5)$, $(5/9, 4/9)$ find prob. that he will be happy



$$P(H) = \frac{2}{5} \times \frac{3}{5} + \frac{1}{5} \times \frac{1}{5} + \frac{2}{5} \times \frac{5}{9}$$

Now, we know that he is happy, find the prob that he is happy since he got the job

Since we know He is happy so we will remove sad part from sample space i.e our sample space has been reduced "Happy" only since we already know he is happy.

So sample space will be

$S = \{(IIIT \,\&\, H), (IIT \,\&\, H), (Job \,\&\, H)\}$

& we want to find the prob. that he is happy when he got the job will be

$E = \{(Job \,\&\, H)\}$ ← Must be subset of "S"

$$\therefore\ P(E) = \sum_{x_i \in E} P(x_i) = \frac{\frac{2}{5} \times \frac{5}{9}}{\frac{2}{5} \times \frac{3}{5} + \frac{1}{5} \times \frac{1}{5} + \frac{2}{5} \times \frac{5}{9}}$$

Note: Sample space can be

1) Finite sample space
2) Countable infinite sample space
3) Uncountable sample space