60004160117 (1)

Probabilistic Analysis (It uses probability for the analysis of algorithms)

*] Hiring problem :-

→ Manger wants to hire an assistance
→ Employment agency

strategy of agency

→ they will send one candidate each day
→ if current assistant, is low
means candidate is better than current assistant
pay to current assistant
hire new candidate

Algorithm :-

```
Hire-assistant (n)
1: best = 0        // candidate 0 is a least-qualified
                      dummy candidate.
2: for i=1 to n
3:   interview candidat i
4:   if candidate i is better than candidate best
5:        best = i
6:        hire candidate i
```

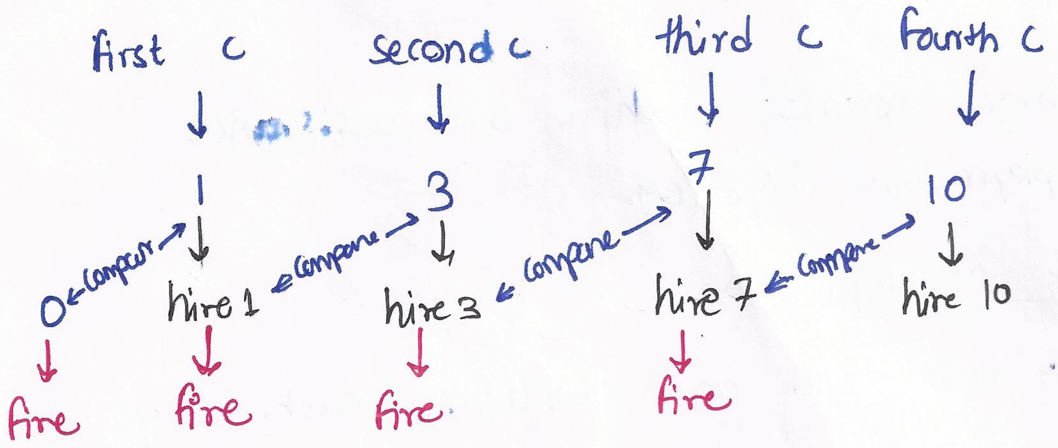
we are not botthered about the time but botthered about the cost

*] low cost - interviewing candidate
c_i

*] Expensive → some money to current assistant
→ agency

worst case :-

① if the candidates are in incrementing order of quality



So this is worst case where we end up hiring

$O(c_h\, n)$

Because we dont have the control on order in which they appear for interview

Here, probabilistic analysis will help to put a bit control on order to get better analysis.

→ In order to use probabilistic analysis we use knowledge of data and

→ make certain assumptions on distribution of ~~data~~ input

→ then analyse & compute average case running time

→ In hiring problem we are more focused on cost

∴ in hiring problem we give rank

∴ candidate i rank(i)

assumption ① :- candidate appear in random order

∴ ranks will form uniform random permutation

$n!$ permutation with equal probability

*] Randomized algorithm

1. In hiring problem we assume that candidates appear in random order, but what if they are not in random
↓
and agency is purposly sending the candidates is increasing order to get the money

2. Control over order is needed

∴ change the model little bit :-

i) list of 'n' candidates is with ageney.

→ get the list from agency in advance so that the probability of candidates getting sent in increasing order will be less.

ii) Use any random generator function to randomize the list

∴ An algorithms which is randomized in behaviour if its output is not only dependent on the input but also dependend on the input produced by random generator

& when the input is random we refer to the running time as expected running time.

*] Indicator random variable

it is method to convert between probabilites & expectations

indicator variable $I\{A\}$

$$I\{A\} = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{if not occurs} \end{cases}$$

$$E[X_h] = E[I\{H\}]$$
$$= 1 \cdot Pr\{H\} + 0 \cdot Pr\{T\}$$
$$= 1 \times \tfrac{1}{2} + 0 \times \tfrac{1}{2}$$
$$= \tfrac{1}{2} \longrightarrow (*)$$

$\therefore$ <u>Lemma</u> :- given S and event A

let $X_A = I\{A\}$

$$\boxed{E[X_A] = Pr\{A\}} \leftarrow \text{for one flip}$$

for n flips

$$X = \sum_{i=1}^{n} X_i$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] \quad \text{solve by linearity}$$
$$= \sum_{i=1}^{n} E[x_i]$$
$$= \sum_{i=1}^{n} \tfrac{1}{2} \longleftarrow x_i = \tfrac{1}{2} \text{ by eq}^n (*)$$
$$= \tfrac{1}{2} + \tfrac{1}{2} \cdots\cdots$$
$$= \tfrac{n}{2}$$

Aanalysis of hiring problem using IRV

let x be random variables = no. of times we hire a new assistant

x_i be indicator random variable for i^{th} candidate

$$x_i = \begin{cases} 1 & \text{if hired} \\ 0 & \text{if not hired} \end{cases}$$

to get hired, candidate i should be better for 1 to $i-1$

$\therefore$ candidate i will get hired is $1/i$

as the order is random & quality is equally likely

$$i = \left(\frac{1}{i}\right)$$

$$\therefore E[X_i] = \frac{1}{i}$$

$$E(X) = E\left[\sum_{i=1}^{n} x_i\right]$$

$$= \sum_{i=1}^{n} E[x_i]$$

$$= \sum_{i=1}^{n} \frac{1}{i}$$

$$= \lg n + O(1)$$

for +ve integers

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} \ldots$$

$$= \sum_{i=1}^{n} \frac{1}{k} = \lg n + O(1)$$

n^{th} harmonic

$\therefore$ we interviewed $\underline{n}$ people

& hired $\underline{\ln n}$

a $\boxed{O(C_h \ln n)}$

IRV

$\boxed{O(C_h n)}$

The birthday paradox

Problem:- How many people must be there in a room to have same birthday

assume there are n days in a year

$n = 365$

$1, 2, \dots k$ people in room

So, let b_i is the birthday of ith person

$$1 \le b_i \le n$$

assuming birthay uniformly distributed

$$\Pr\{b_i = r\} = \frac{1}{n} \quad \text{for } i = 1 \dots k, \quad r = 1 \dots n$$

→ random selection of birthdays are independent probability ith person birthay & jth person birthday will fall on day r

$$\Pr\{b_i = r \text{ and } b_j = r\} = \Pr\{b_i = r\} \cdot \Pr\{b_j = r\}$$

$$= \frac{1}{n} \times \frac{1}{n}$$

$$= \frac{1}{n^2}$$

∴ Both will fall on same day

$$\Pr\{b_i = b_j\} = \sum_{r=1}^{n} \Pr[b_i = r \;\&\; b_j = r]$$

$$= \sum_{r=1}^{n} \left(\frac{1}{n^2}\right)$$

Using ~~IVR~~ IRV

For pair $\{i, j\}$ of the k people in a room

x_{ij} is indicator variable or
Indicator random variable

$$1 \le i < j < k$$

$$x_{ij} = \begin{cases} 1 & \text{if } i \,\&\, j \text{ have same birthday} \\ 0 & \text{otherwise} \end{cases}$$

$$E[X_{ij}] = Pr[\text{Person } i \,\&\, j \text{ same birthday}]$$
$$= \frac{1}{n}$$

x be the random variable that counts number of pair

$$x = \sum_{i=1}^{k} \sum_{j=i+1}^{k} x_{ij}$$

taking expectation on both sides

$$E[x] = E\left[\sum_{i=1}^{k} \sum_{j=i+1}^{k} x_{ij}\right] \quad \text{apply linearity}$$

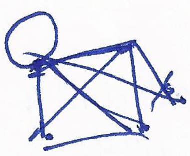
$$= \sum_{i=1}^{k} \sum_{j=i+1}^{k} E[x_{ij}]$$

$$= \frac{k(k-1)}{2n} \ge 1$$

$$\boxed{k(k-1) \ge 2n}$$

$n(n-1)$

n n^2 n^3



math

$$\boxed{k = \sqrt{2n} + 1}$$

$k = 28$

$n = 365$

$$k(k-1) \quad \frac{(28)(27)}{} \approx 1.035$$

Balls & Bins

bin model & # model are same

let we have b bins

we randomly toss balls into these bins,

probability equally likely

Probability the ball land in given bin $= \frac{1}{b}$

① How many balls will fall in a given bin ?

$\boxed{\frac{n}{b}} \longleftrightarrow$ {Binomial distribution}

according to

② How many balls we must toss on average until a bin contains a ball

expected success $= \frac{1}{\text{probability of Success}} = \frac{1}{1/b} = \underline{\underline{b}}$

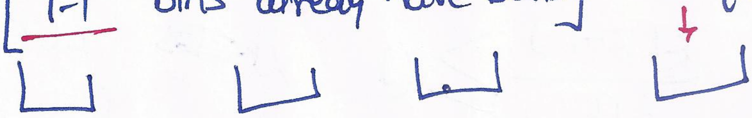
③ How many balls must be toss so, empty bin contains at least one ball.

Say when a ball hands in a bin we call it hit

bin $\longrightarrow$ "hit"

i th stage when $(i-1)$ hits have occured

[i-1 bins alreay have balls] ith



for each toss in the ith stage, {as $i-1$ bins contains balls}

$b-(i-1)$ are empty $\longrightarrow$ Expand this $b-i+1$

probability of hit at ith stage

$$p = \frac{b-i+1}{b}$$

Expectation of geometric distribution

$$E(x) = \frac{1}{p}$$

for
$$E(n_i) = \frac{1}{\frac{b-i+1}{b}}$$

$$= \frac{b}{b-i+1}$$

for n tosses

$$E(n) = E\left[\sum_{i=1}^{b} n_i\right]$$ following linearity

$$= \sum_{i=1}^{b} E(n_i)$$

$$= \sum_{i=1}^{b} \frac{b}{b-i+1}$$

if put values of i

$$= \frac{1}{b-1+1} + \frac{1}{b-2+1} + \dots \frac{1}{b}$$

$$= \sum_{i=1}^{b} \frac{1}{b\,i}$$

$$= b \sum_{i=1}^{b} \frac{1}{i}$$

$$= b(\ln b + O(1))$$ using nth harmonic