SD Curs 3

12 Mar 2025

1. Variabile aleatoare
    - Moneda
    - Zaruri
    - Secretary Problem
    - Birthday Paradox
2. Quicksort

- $x = \begin{cases} 0 & \text{dacă} \quad \text{Heads (cap)} \\ 1 & \text{dacă} \quad \text{Tails (pajură)} \end{cases}$

$Pr(x = 0) = \frac{1}{2}$

$Pr(x = 1) = \frac{1}{2}$

$\mathbb{E}[x] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$

medie
„expectation"

- x, dat cu zarul

$x = \begin{cases} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{cases}$

$Pr(x = i) = \frac{1}{6} \quad \forall \; 1 \le i \le 6$

$\mathbb{E}[x] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots + \frac{1}{6} \cdot 6$

$= \frac{7 \cdot 6}{2} \cdot \frac{1}{6} = \frac{7}{2} = 3,5$

- Variabile aleatoare care au valoarea 0 sau 1

$x = \begin{cases} 0 \\ 1 \end{cases}$

$\mathbb{E}[x] = Pr(x)$

ex The secretary problem

Avem n candidați la un post care vin în ordine aleatorie

Intervievăm pers. în ordine și dacă persoana i este mai bună decât primele $i-1$ persoane, atunci angajam persoana i și plătim x \$ compensație

<u>Întrebare</u>. Care este numărul mediu de persoane angajate?

Exemplu

$\underline{2} \quad 1 \quad \underline{3} \quad \underline{4} \quad \underline{6} \quad 5$

<u>Sol</u>:

$$x_i = \begin{cases} 0 & \text{dacă pers. } i \text{ nu este angajată} \\ 1 & \text{dacă pers. } i \text{ este angajată} \end{cases}$$

$$x = x_1 + x_2 + \dots + x_n$$

$$\mathbb{E}[x] = ?$$

$$\mathbb{E}[x] = \mathbb{E}[x_1 + x_2 + \dots + x_n] = \mathbb{E}[x_1] + \dots + \mathbb{E}[x_n]$$

$$\mathbb{E}[x_i] = P(x_i = 1) = \frac{1}{i}$$ probabilitatea ca candidatul i să fie mai bun decât primii $i-1$

$$\mathbb{E}[x] = \sum_{i=1}^{n} \frac{1}{i} = O(\log n)$$

$\square$

ex Birthday Paradox

n studenți

Câți studenți (în medie) aceeași zi de naștere

Sol:

$$x_{ij} = \begin{cases} 1 & \text{dacă studentul } i \text{ și studentul } j \\ & \text{au aceeași zi de naștere} \\ 0 & \text{altfel} \end{cases}$$

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij}$$

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij} \right]$$

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

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(x_{ij} = 1)$$

$$P(x_{ij} = 1) = \sum_{k=1}^{365} P\left(\begin{array}{c} \text{st } i \text{ să fie născut în ziua } k \\ \& \\ \text{st } j \text{ să fie născut în ziua } k \end{array} \right)$$

$$= \sum_{k=1}^{365} P(\text{st } i \text{ să fie născut în ziua } k) \cdot P(\text{st } j \text{ să fie născut în ziua } k)$$

$$= \sum_{k=1}^{365} \frac{1}{365^2} = 365 \cdot \frac{1}{365^2} = \frac{1}{365}$$

$$\mathbb{E}[X] = \frac{n(n-1)}{2 \cdot 365} \qquad \square$$

Quicksort

```
Quicksort ( A, l, r )
{
    q = Partition ( A, l, r )
    Quicksort ( A, l, q-1 )
    Quicksort ( A, q+1, r )
}
```

3 7 2 11 22 9 -4 $\boxed{5}$

pivot 5 -4 3 2 $\boxed{5}$ 7 11 22 9

$q = 4$

pivot 2 -4 2 3

---

aşezarea pivotului

$$T(n) = 2\,T\left(\frac{n}{2}\right) + O(n) = O(n \log n)$$

în cazul cel mai favorabil

Pivotul pe ultima poziție

1 2 3 4 . $\square$ $\boxed{n}$

$$T(n) = T(n-1) + O(n) = O(n^2)$$

Notăm cu z_1 cel mai mic el. din vector, z_2 al doilea cel mai mic etc

Vom defini valoarea aleatoare X

$$X_{ij} = \begin{cases} 1 & \text{daca } z_i \text{ și } z_j \text{ sunt comparate la un pas al algoritmului} \\ 0 & \text{altfel} \end{cases}$$

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij} \leftarrow \text{numarul de comparații}$$

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} x_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathbb{E}[x_{ij}]$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathcal{P}(x_{ij} = 1)$$

$z_1 \; z_2 \; \cdots \; z_i \; z_{i+1} \; \cdots \; z_j \; z_{j+1} \; \cdots \; z_n$

$$\mathcal{P}(x_{ij} = 1) = \frac{2}{j-1+i}$$

2	7	5	-4	22	10	8	3	1
z_3	z_6	z_5	z_1	z_9	z_8	z_7	z_4	z_2

Dacă alegem 7 (z_6) ca pivot, z_5 și z_7 nu se compară

2 <u>5</u> -4 1 3 $\boxed{7}$ 8 <u>10</u> 22

<u>Obs</u> Numerele consecutive se pot compara mereu

$$\mathbb{E}[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1}$$

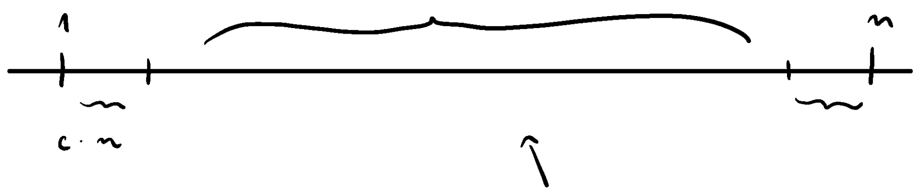
$$= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1} < \sum_{i=1}^{n} \underbrace{\sum_{k=1}^{n} \frac{2}{k}}_{\log n} = \sum_{i=1}^{n} O(\log n)$$

$$= O(n \log n)$$

s.v. $k = j - i$

$$T(n) = T\left(\frac{n}{3}\right) + T\left(\frac{2n}{3}\right) = O(n \log n)$$

$$T(n) = T\left(\frac{n}{10000}\right) + T\left(\frac{9999\,n}{10000}\right) + O(n) = O(n \log n)$$



c = constantă