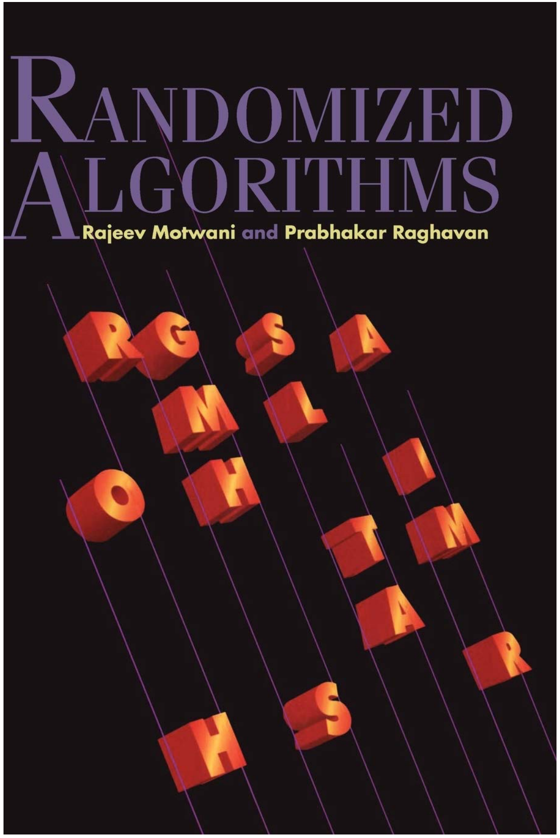
COL 758: Advanced Algorithms

Lecture 08

Reference Book



Randomized Algorithms:
Motwani, Raghavan

Puzzle: Estimating Set Size

There are p players playing the following game:

Game: Each player spins a wheel to generate a random number in $[0,1]$, and winner is the one with least score.

Rounds: k

Available information: Winning scores in all k rounds.

How to efficiently predict "p"?

Winning Scores
0.1
0.12
0.3

  . . . . 

0.2	0.1		0.4
0.42	0.3		0.12
0.3	0.85		0.6

Puzzle: Estimating Set Size

There are p players playing the following game:

Game: Each player spins a wheel to generate a random number in $[0,1]$, and winner is the one with least score.

Rounds: k

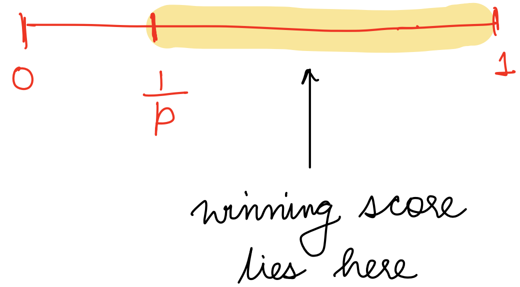
Available information: Winning scores in all k rounds.

How to efficiently predict "p"?

The expected value of winning score in each round is $\dfrac{1}{p+1}$

Question 1

If p points are sampled uniformly randomly from $[0,1]$, then what is probability that all of them are larger than $1/p$?

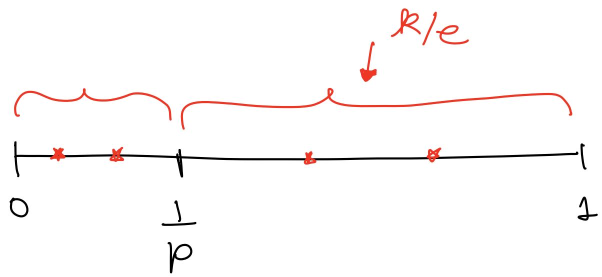


Solution:

$$\left(1 - \frac{1}{p}\right)^p \approx \frac{1}{e}$$

Question 2

If $w_1, w_2, \ldots, w_k$ are winning scores of k rounds in the ***game***, then on expectation how many of them are larger than $1/p$?



Solution:

$$\frac{k}{e}$$

Weak Predictor

$$\frac{1}{\left(\sum_{i=1}^{k} w_i\right)/k}$$

σ^2 variance of (w_1) variance $\left(\sum_{i=1}^{k} w_i/k\right)$ $\frac{\sigma^2}{k}$

differ by factor of k

$k \approx \log(p)$

There are p players playing the following game:

Game: Each player spins a wheel to generate a random number in $[0,1]$, and winner the is one with least score.

Rounds: k

Available information: Winning scores in all k rounds.

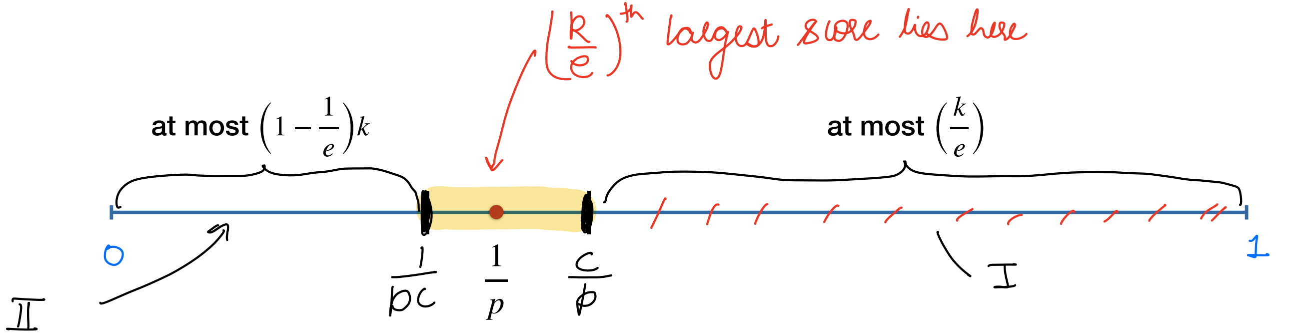
How to efficiently predict "p"?

Algorithm

1. Let $w_1, w_2, \ldots, w_k$ be winning scores of k rounds in the ***game.***
2. w_0 = $(k/e)^{th}$ largest score.
3. Output $\dfrac{1}{w_0}$

Time $= O(k \log k)$

$\boxed{e^{-\frac{\mu\delta^2}{2+\delta}}}$



Ques: What is probability of event that more than (k/e) winning scores are larger than $\frac{c}{p}$?

Sol^n: For $i=1$ to k, $X_i = \begin{cases} 1 & \text{score } w_i > c/p \\ 0 & \text{o/w} \end{cases}$

$$E[X_i] = \text{Prob}[X_i = 1] = \left(1 - \frac{c}{p}\right)^p = \left(1 - \frac{c}{p}\right)^{\frac{p}{c} * c} \approx \frac{1}{e^c} \leq \frac{1}{2e} \quad \text{Take } c = 1.7$$

$X = X_1 + X_2 + \dots + X_k$ = no of scores larger than c/p.

$$\boxed{E[X] \leq \frac{k}{2e}}$$

$$\text{Prob}\left[X > \frac{k}{e}\right] \leq \text{Prob}\left[X > \underset{\delta = 1}{2E[X]}\right] \leq e^{-\frac{E[X]}{3}} \leq e^{-\frac{k}{e^c \cdot 3}} = \frac{1}{p^3}$$

$k = e^c \cdot 9 \cdot \log(p)$

Approximate Transitive closure

Given: A directed graph $G=(V,E)$ with n *vertices*, m edges

Find: Estimate size of transitive closure, for each vertex v *in* V.

If $p(v)$ = no. of vertices reachable from v, then

the estimate $\hat{p}(v)$ must lie in $\left[\frac{p(v)}{c}, c \cdot p(v)\right]$

Approach 1 : BFS/DFS from all vertices − $O(mn)$

Approach 2 : Adj matrix "A", repeated matrix product − $\Omega(n^{\omega})$

Current Best Det Algo.

$\Omega(\min\{m \cdot n, n^{\omega}\})$

Approximate Transitive closure

Given: A directed graph $G=(V,E)$ with n *vertices*.

Find: Estimate size of transitive closure, for each vertex v *in* V.

If $p(v)$ = no. of vertices reachable from v, then

the estimate $\hat{p}(v)$ must lie in $\left[\frac{p(v)}{c},\ c \cdot p(v)\right]$

Algorithm

For ($i = 1$ to k) do:

- Assign to each $v \in V$ a uniformly independent random number from [0,1].
- For each ($v \in V$): $w_i[v]$ = Label of smallest label vertex reachable from v.

For each ($v \in V$):

- let $w[v] = (k/e)^{th}$ largest element from the set $\{w_1[v], \ldots, w_k[v]\}$.
- Output $\hat{p}(v) = \dfrac{1}{w[v]}$