EXERCISE 3 Consider the following algorithm:

SELECT(S, i) $\left\{ \begin{array}{l} 1 \le i \le |S| \\ \text{selects } i\text{-th o.s.} \end{array} \right\}$

$n \leftarrow |S|$

(B) if $(n=1)$ then return $S[1]$

(D) $p \leftarrow$ RANDOM(S)

$S_1 \leftarrow \{s \in S : s < p\}$; $S_2 \leftarrow \{s \in S : s > p\}$

if $(|S_1| = i-1)$ then return p

(R+C) if $(|S_1| > i-1)$

then return SELECT(S_1, i)

else return SELECT$(S_2, i - |S_1| - 1)$

1. Prove that SELECT(S, i) correctly returns the i-th order-statistic of S (i.e., i-th smallest element)

2. Analyze its average running time

REMARK

- Unlike QUICKSORT, we cannot prove that

$$T_{SELECT}(n) = \Theta(n) \text{ w.h.p.!}$$

- High-probability randomized selection is possible:

IDEA: HP_SELECT(S, i)

- based on large sample selection

$\Theta(n^{3/4})$ — pick $\sim n^{3/4}$ random elements

$\Theta(n^{3/4}\log n^{3/4}) = \Theta(n^{3/4}\log n)$ — sort them

$\Theta(1)$ — pick $\underline{a, b}$ of rank $\sim i \cdot n^{-1/4} \pm \sqrt{n}$ in the sequence

$\Longrightarrow$

$\Theta(n)$ w.h.p. $\left\{\begin{array}{l} \text{Set } S' = \{s \in S : a \leq s \leq b\} \\ \text{1) contains } i\text{-th o.s.} \\ \text{2) has size } O(n^{3/4}) \end{array}\right.$

$\Longrightarrow$ Sort S' to identify i-th o.s.

$\Theta(n^{3/4}\log n)$

EXERCISE 4 Consider k i.i.d. geometric variables:

$$Z_1, \dots, Z_k \sim \mathrm{Geom}(p) \qquad 0<p<1$$

$$\forall i : \Pr(Z_i = j) = p(1-p)^{j-1}, \quad j \in \mathbb{Z}^+$$

Recall that $E[Z_i] = \frac{1}{p}$.

Let $X = \sum_{i=1}^{k} Z_i$.

Then $\mu = E[X] = \sum_{i=1}^{k} E[Z_i] = \frac{k}{p}$.

Discuss how to use Chernoff's bounds to upper bound

$$\Pr(X > t\mu) \qquad t \in \mathbb{Z}^+$$

(Hint: associate the event to Bernoulli trials)

EXERCISE: Given a sorted array $A[1..n]$, assume that it can be read only via method

$$A.\text{read}(i) = \begin{cases} A[i] & p = 1/2 \\ \text{error} & p = 1/2 \end{cases}$$

(models faulty storage).

Write a randomized binary search routine SEARCH(A, 1, n, K) returning 1 $\iff \exists i : A[i] = K$ executing $O(\log n)$ reads w.h.p.

HINT: Same structure of binary search but each value is read multiple times until A.read(·) ≠ error. Use the geometric bound in the analysis.

EXERCISE 4 Consider n i.i.d. variables $X_i \in \{-1, 1\}$, with

$$\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$$

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

1. Show how to adapt Chernoff Bound 2 to get an upper bound
$$\Pr(X > \delta), \quad \delta > 0$$

2. Determine a value δ_n : $\Pr(X > \delta_n) < \frac{1}{n}$

EXERCISE 5: Consider the unit square $S := [0,1] \times [0,1] \in \mathbb{R}^2$ and its K^2 subsquares of size $1/K \times 1/K$. How many calls to RANDOM(S) to cover all subsquares with $p \geq 1 - \frac{1}{K^2}$?