

EXERCISE 3 Consider the following algorithm:

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

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

(D) $P \leftarrow \text{RANDOM}(S)$

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

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

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

then return $\text{SELECT}(S_1, i)$

else return $\text{SELECT}(S_2, i - |S_1| - 1)$

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

2. Analyze its average running time

1. induzione su n :

- base: $n=1 \Rightarrow$ OK

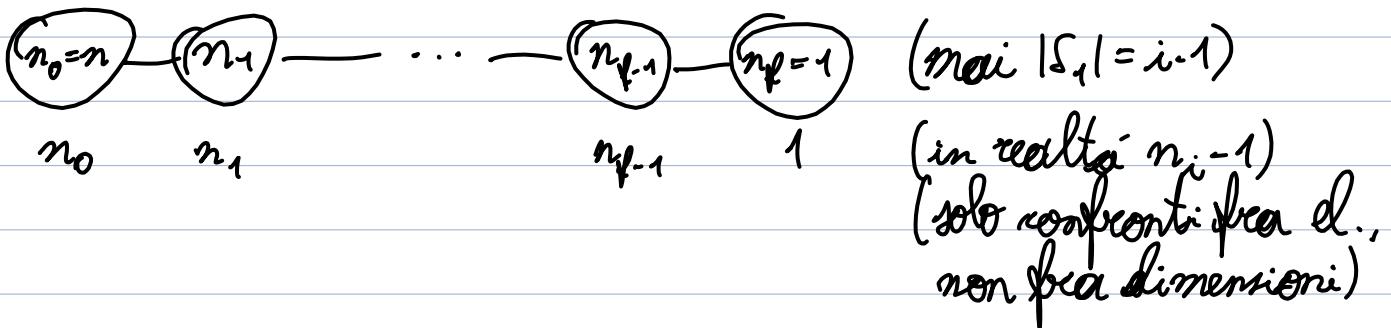
- induzione: $n > 1$:

□ $|S_1| = i-1 \Rightarrow S_1$ contiene $x_1, \dots, x_{i-1} \Rightarrow$ subito dopo c'è $x_i = p$

□ $|S_1| > i-1 \Rightarrow S_1$ contiene $x_1, \dots, x_{|S_1|} \Rightarrow x_i = p$ sarà in S_1 con range i

□ $|S_1| < i-1 \Rightarrow x_i$ sarà in $S_2 \Rightarrow$ range di x_i in S_2 : $i - |S_1| - 1$

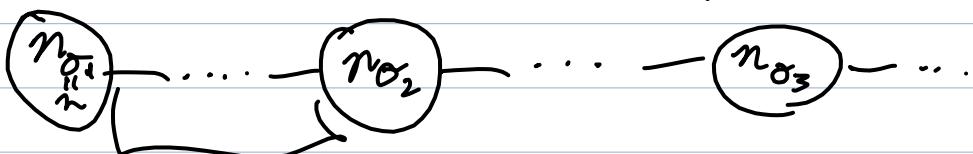
2. w.h.p. daí $O(n \log n) \Rightarrow$ vogliamo tempo lineare
recursion path in worst case:



n_i : r.v. \Rightarrow stesso argomento di quicksort

$$\Pr[\text{scelta pivot fortunata}] = \Pr[n_i \leq \frac{3}{4} n_{i-1}] \geq 1/2$$

Dividiamo recursion path in segmenti $\alpha_1, \alpha_2, \dots$



qui ho ridotto

dimensione di $3/4 \Rightarrow n_{\alpha_i} \leq \frac{3}{4} n_{\alpha_{i-1}}$

k^* : ultimo segmento è $n_{\alpha_{k^*}} \Rightarrow k^* \leq \lceil \log_{3/4} n \rceil + 1 \left(\Rightarrow n_{\alpha_{k^*}} \leq \left(\frac{3}{4}\right)^{k^*} n \right)$

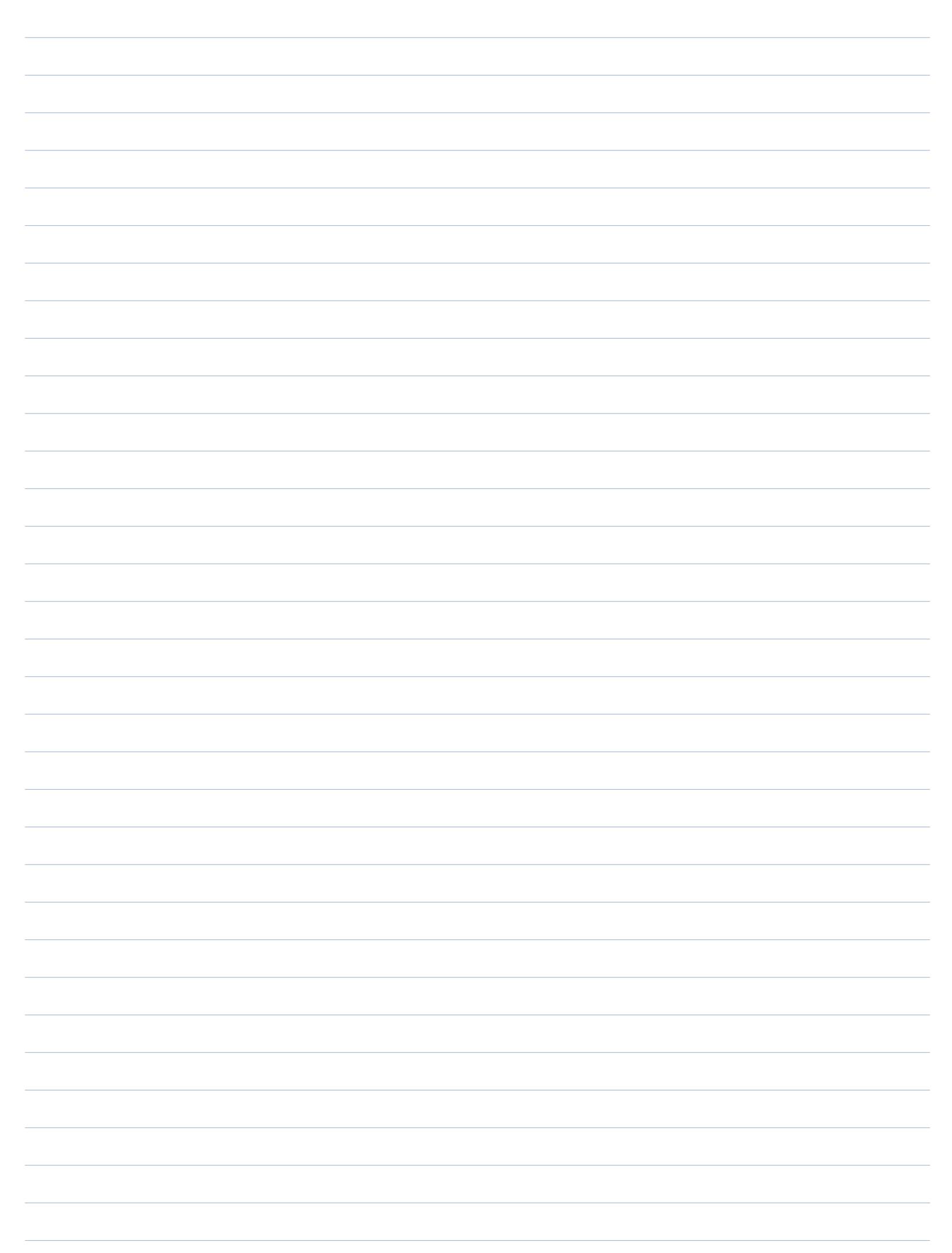
Tutti nodi in segmento α_i fanno lavoro $\leq n_{\alpha_i}$

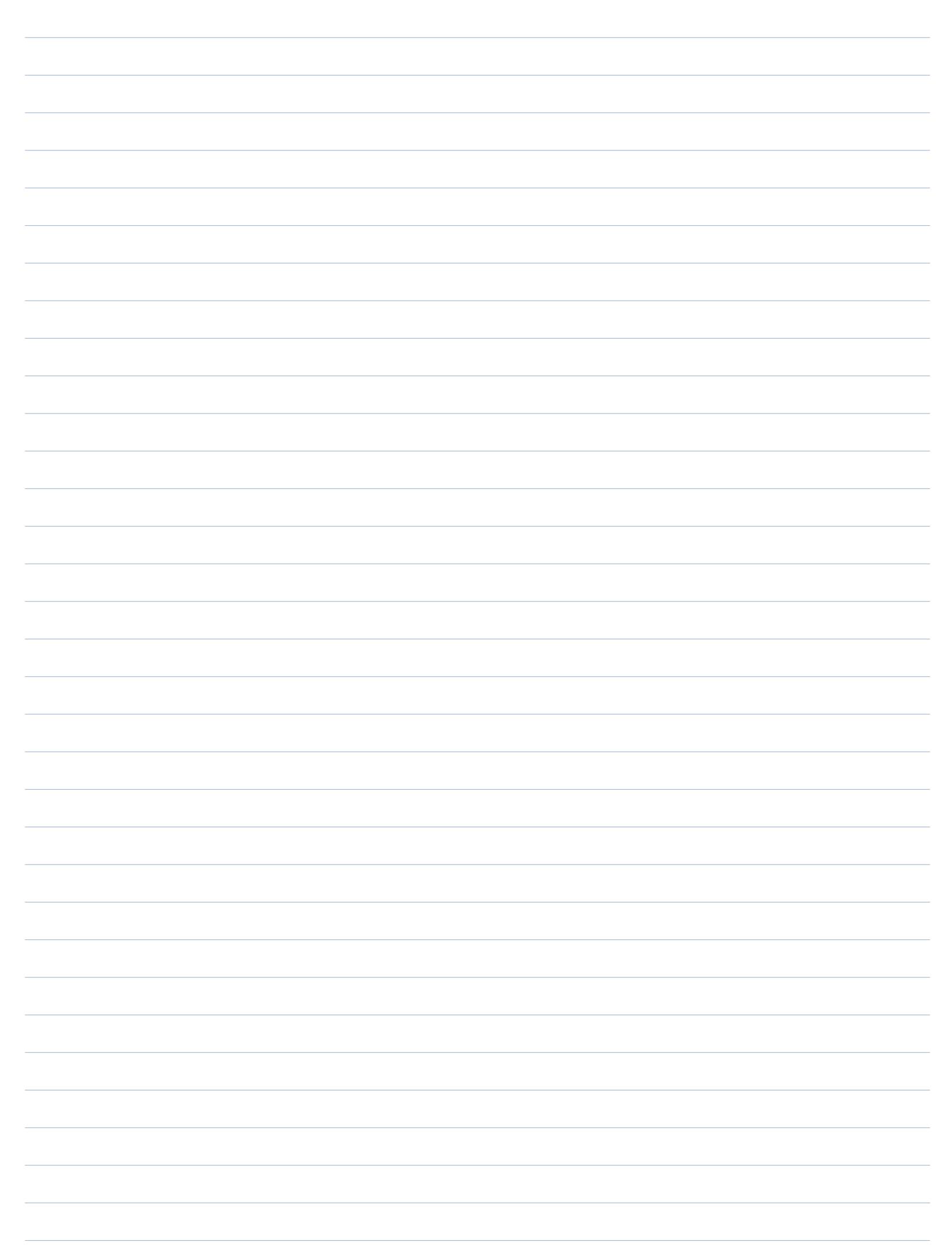
$$\begin{aligned} \mathbb{E}[T(n)] &= \mathbb{E}\left[\sum_{i=1}^{k^*} n_{\alpha_i} \cdot |\alpha_i|\right] = \mathbb{E}\left[\sum_{i=1}^{k^*} \left(\frac{3}{4}\right)^{i-1} n \cdot |\alpha_i|\right] = \\ &= \sum_{i=1}^{k^*} \mathbb{E}\left[\left(\frac{3}{4}\right)^{i-1} n \mid |\alpha_i|\right] = \sum_{i=1}^{k^*} \left(\frac{3}{4}\right)^{i-1} n \mathbb{E}[|\alpha_i|] \end{aligned}$$

$p \geq 1/2$: scelta fortunata \Rightarrow avanti in segmento finché non fa prenbo
 $\Rightarrow |\alpha_i| \sim \text{geom}(p \geq 1/2) \Rightarrow \mathbb{E}[|\alpha_i|] = 1/p \leq 2 \Leftrightarrow$

$$\Rightarrow \mathbb{E}[T(n)] \leq \left(\sum_{i=1}^{k^*} \left(\frac{3}{4}\right)^{i-1}\right) 2n \leq \left(\sum_{i=1}^{\infty} \left(\frac{3}{4}\right)^{i-1}\right) 2n = \left(\sum_{j=0}^{\infty} \left(\frac{3}{4}\right)^j\right) 2n = \frac{1}{1 - \frac{3}{4}} 2n$$

W.h.p.: $c \log n$ nodi \Rightarrow in questi livelli non sappiamo dove
avremo scelta fortunata \Rightarrow contributi diverse





REMARK

- Unlike **QUICKSORT**, we cannot prove that

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

- High-probability randomized selection is possible :

IDEA : HP. $\text{SELECT}(S, i)$

- based on large sample

$\Theta(n^{3/4})$ selection
 $\Theta(n^{3/4} \log n)$
 $= \Theta(n^{3/4} \log n)$
 $\Theta(1)$

pick $n^{3/4}$ random elements
 sort them
 pick a, b of rank
 $n^{11/4} + \sqrt{n}$ in the sequence

\Rightarrow

$\Theta(n)$
 w.h.p.

Set $S' = \{s \in S : \text{ess} \leq b\}$

1) contains i -th O.S.
 2) has size $\Theta(n^{3/4})$

\Rightarrow Sort S' to identify i -th O.S.
 $\Theta(n^{3/4} \log n)$

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

$$Z_1, \dots, Z_k \sim \text{geom}(p) \quad 0 < p < 1$$

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

Recall that $E[Z_i] = \frac{1}{p}$.
Let $X = \sum_{i=1}^k Z_i$.

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

Discuss how to use Chernoff's bounds to upper bound

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

(Hint: associate the event to Bernoulli trials)

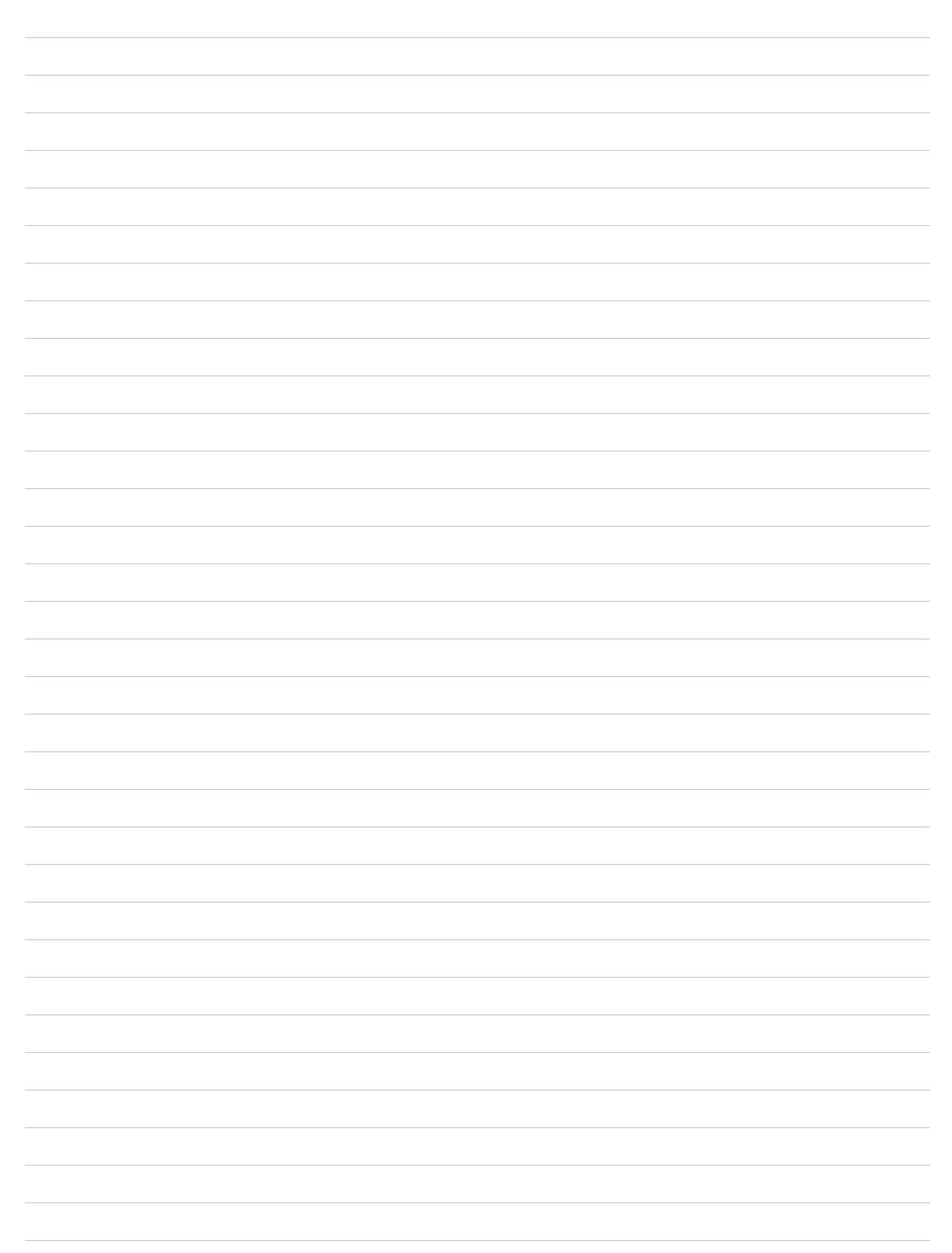
Collegiamo X a sequenza di lanci monete

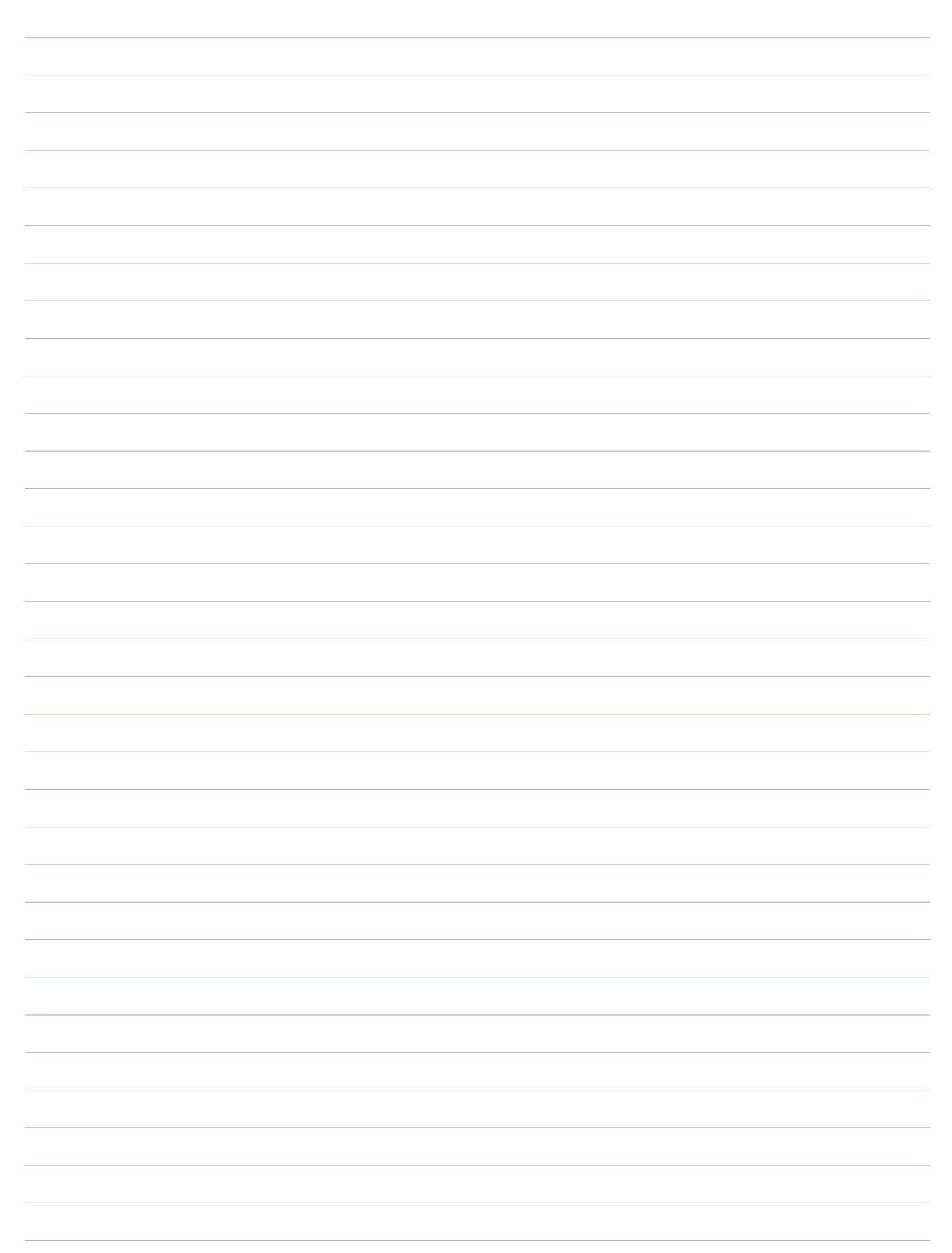
Z_i : i -lancio con i -esimo successo dopo $(i-1)$ -esimo

$$\left. \begin{array}{l} E_1 = "X > t\mu" \\ E_2 = " $\leq k$ successi in $t\mu$ lanci" \end{array} \right\} E_1 \rightarrow E_2$$

$$Y_1, \dots, Y_{t\mu} \sim \text{Bin}(n) \Rightarrow E_2 = "Y = \sum_{i=1}^{t\mu} Y_i \leq k" \Rightarrow E[Y] = t\mu np = t\frac{k}{p}p = tk \Rightarrow$$

$$\Rightarrow \Pr[E_2] = \Pr[Y \leq E[Y]/t] \Rightarrow \Pr[Y \leq (1-\varepsilon)E[Y]], \quad 1/t = 1 - \varepsilon$$





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
 $\text{SEARCH}(A, l, m, K)$ returning $1 \leq i \leq n : A[i] = K$
executing $O(\log n)$ reads w.l.o.g.

HINT: Same structure of binary search
but each value is read multiple times
until $A.read(\cdot) \neq \text{error}$. Use the
geometric bound in the analysis.

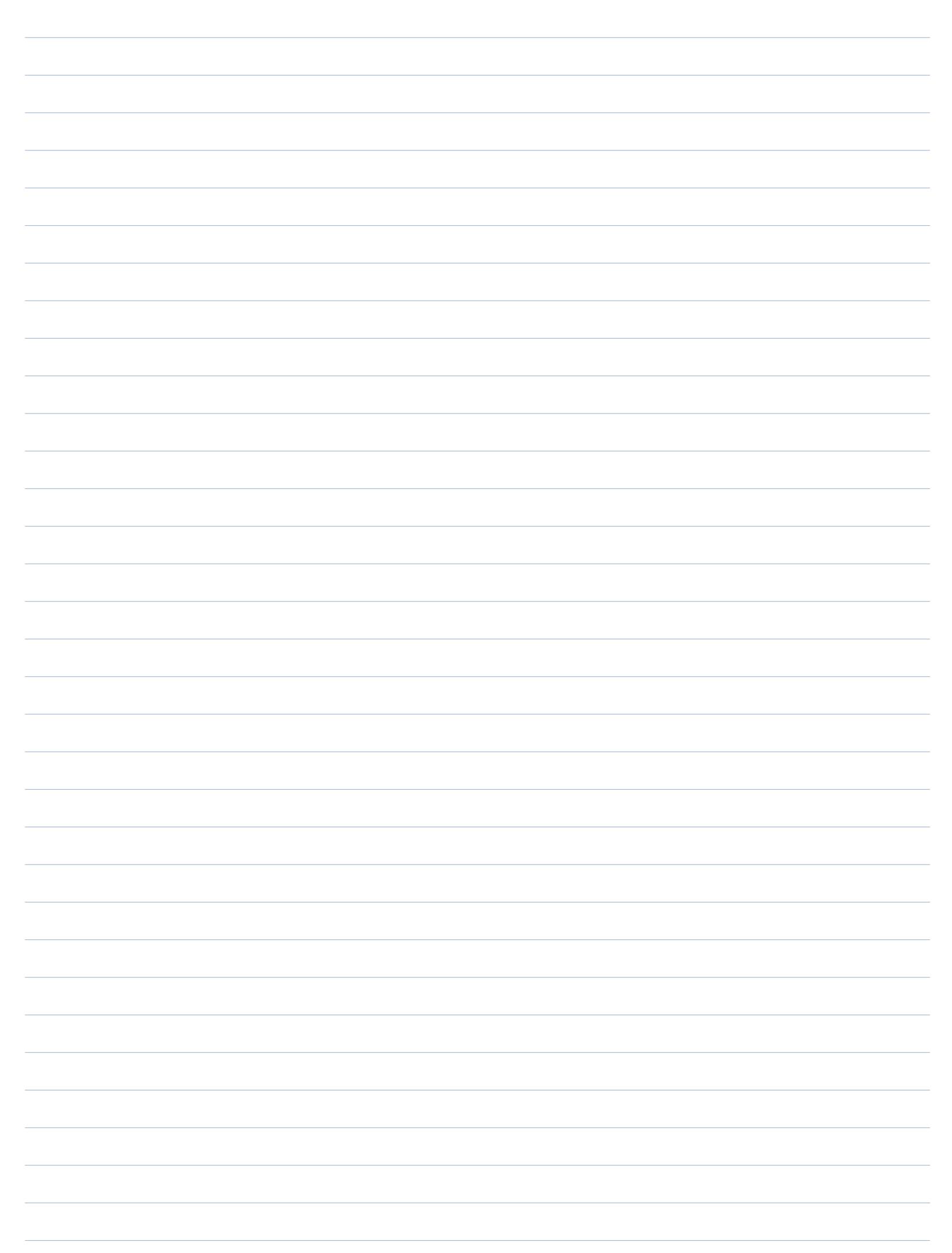
EXERCISE 4 Consider n i.i.d. variables $X_i \in \{-1, 1\}$, with

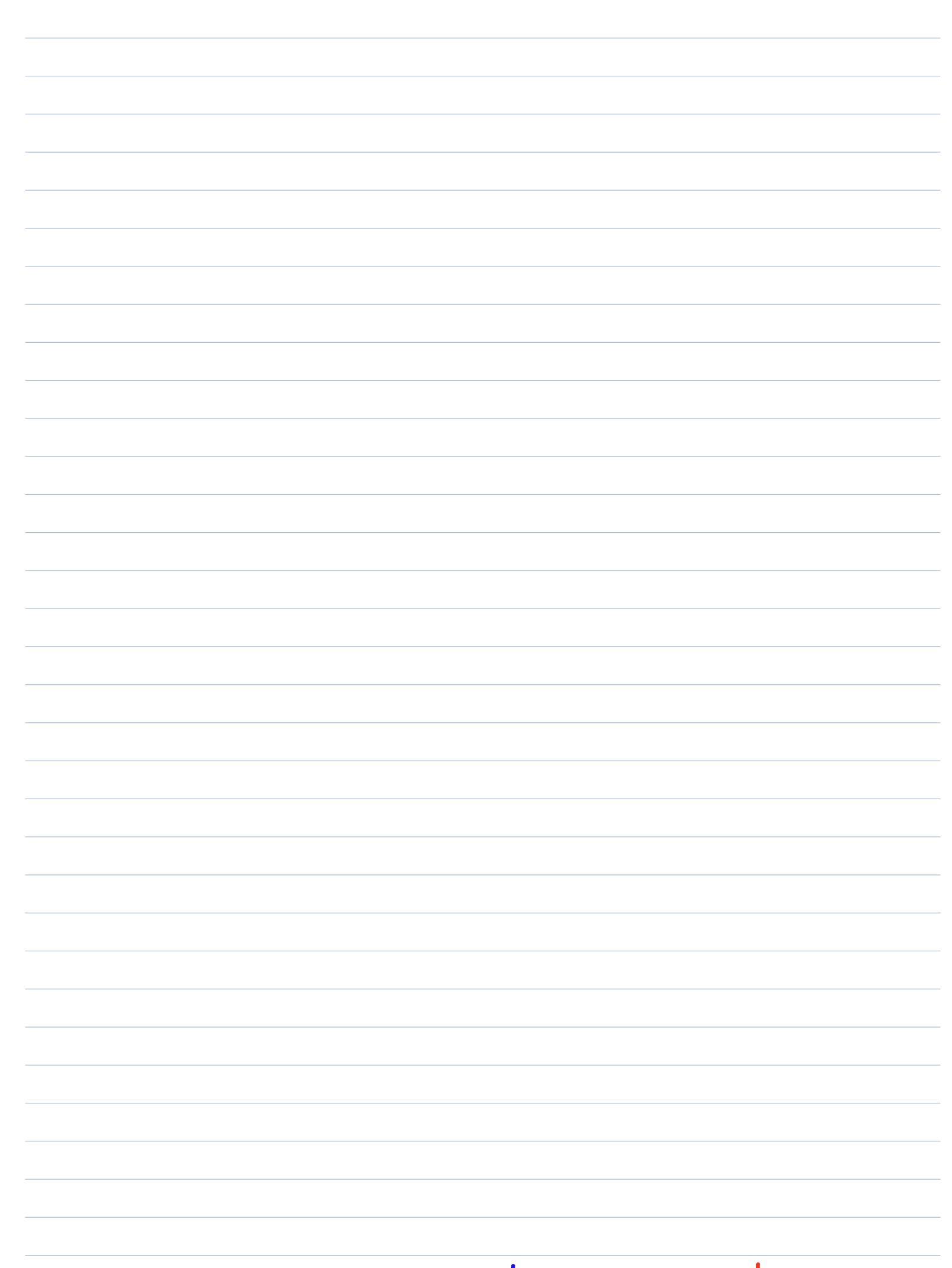
$$\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$$

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

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

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





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 $\text{RANDOM}(S)$ to cover all subsquares with $p \geq 1 - \frac{1}{K^2}$?

