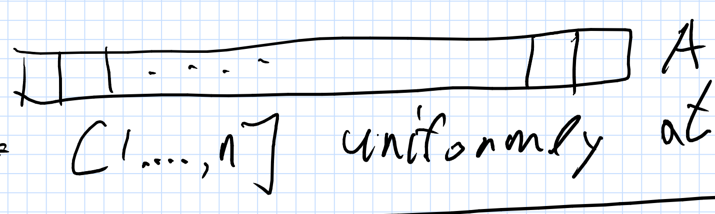
Lecture 5 | 2/9/21

Quick Sort

- Expected # of comparisions.
- High probability

Randomized algorithms

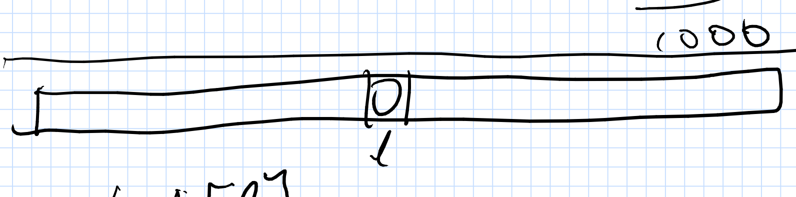


$l \in [1, \dots, n]$ uniformly at random

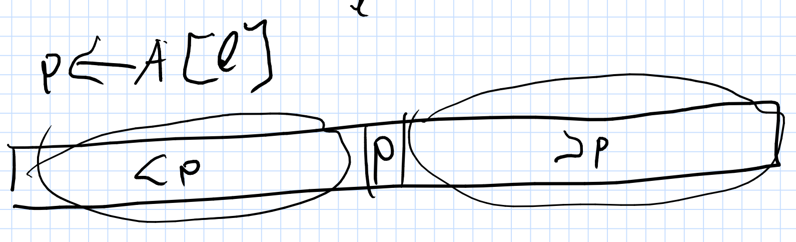
physical experiment | comput.

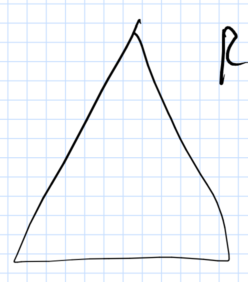
1000

$\frac{1}{1000}$



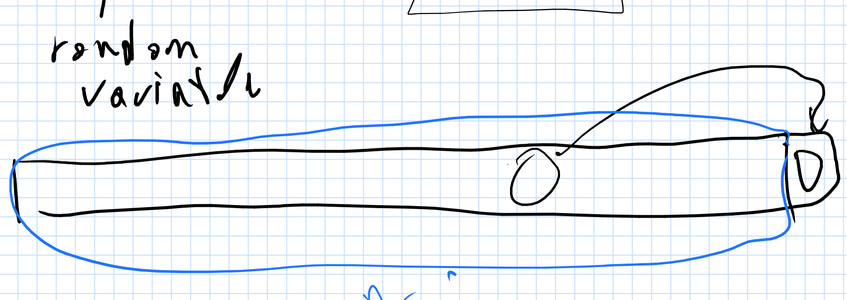
$p \leftarrow A[l]$





Recursion tree execution

RT
↑
random variable



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

X random variable

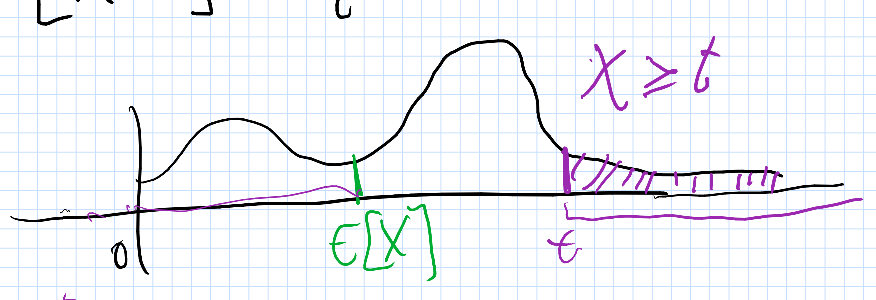
$E[X]$ = average value of $X = \sum_{i=1}^{\infty} P[X=i] \cdot i$

expectation

Markov inequality

Thm
$X \geq 0$ Random variable

$$P[X \geq t] \leq \frac{E[X]}{t}$$



Proof

$P[X \geq t] > \frac{E[X]}{t}$

$\Rightarrow E[X] = \sum i \, P[X=i]$

$\geq \sum_{i \geq t} P[X=i] \cdot i$

$\geq t \sum_{i=t}^{\infty} P[X=i]$

$= t \, P[X \geq t] > t \frac{E[X]}{t} > E[X]$.

Property

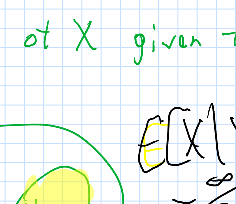
X, Y two RV (not assuming indep).

$E[X+Y] = E[X] + E[Y]$

Linearity of expectation.

Conditional probability

$P[X|Y] = P[X=x \mid Y=y]$



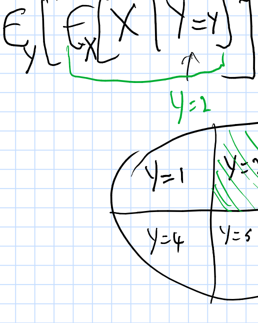
$P[Y=3]$

$P[X=2 \mid Y=1]$

$P[X=x \mid Y=y] = \frac{P[X=x \cap Y=y]}{P[Y=y]}$

$P[X \mid Y=y]$ ← function

$E[X \mid Y=y]$ conditional expectation
||
average of X given that Y=y



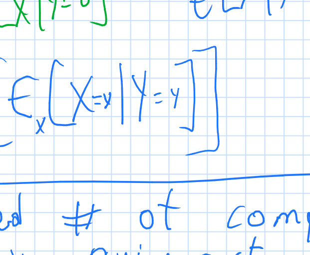
$E[X \mid Y=1] = \sum_{i=1}^{\infty} P[X=i \mid Y=1] \cdot i$

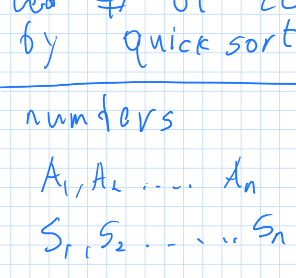
Property conditional expectations

$\forall X, Y$ RV not necessarily independent.

$E[E[X|Y]] = E[X]$ Martingales

$E_y[E_x[X \mid Y=y]] =$





$E[X|Y=0]$ $E[X|Y=1]$

$E_y[E_x[X=x \mid Y=y]]$

Expected # of comparisions performed by quick sort.

n numbers

$A_1, A_2 \dots A_n$ distinct numbers

$S_1, S_2 \dots S_n$ sorted numbers.

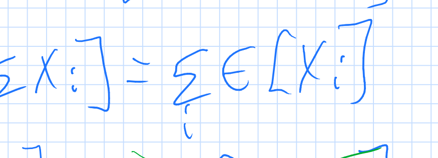
$P[\text{QS compared } S_7 \text{ to } S_8] = 1$

$P[S_i \text{ is compared to } S_j]$

$X_{ij} = 1 \Leftrightarrow S_i$ got compared to S_j

Observations on QS

1. Running time of QS is proportional to the number of comparisions it performs.
2. Two input numbers are compared by QS at most once.



Lemma
The running time of QS is $O\left(\sum_i \sum_{j>i} X_{i,j}\right)$

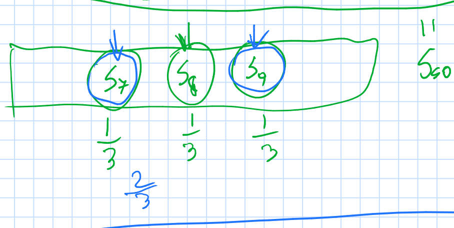
$Y = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{i,j}$

$E[Y] = E\left[\sum_i \sum_j X_{i,j}\right] = \sum_i \sum_j E[X_{i,j}]$

$E[\sum X_i] = \sum_i E[X_i]$

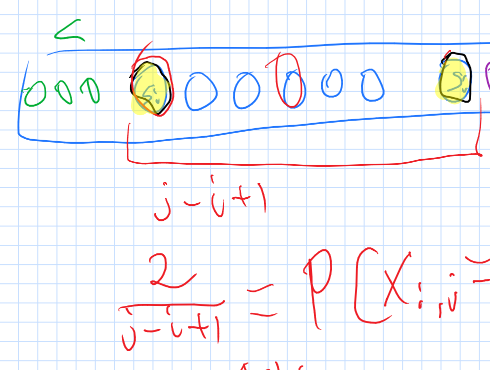
$E[X_{i,j}] = \cancel{0 \cdot P[X_{i,j}=0]} + 1 \cdot P[X_{i,j}=1]$

$= P[X_{i,j}=1]$

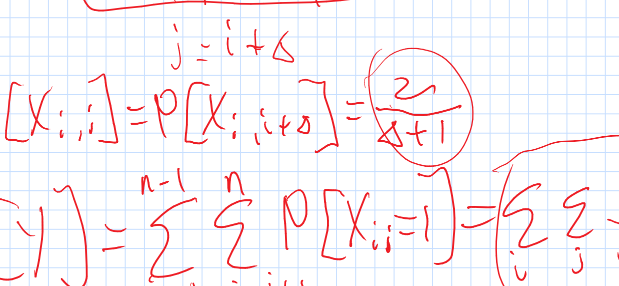


$P[X_{1,n}] = P[\text{QS compares } S_1 \text{ to } S_n \text{ during execution}] = \frac{2}{n}$

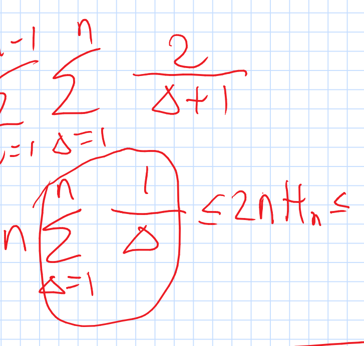
$P[X_{7,9}] = P[S_7 \text{ is compared to } S_9]$



$P[X_{i,j}] = P[\text{QS compares } S_i \text{ to } S_j]$



$\frac{2}{j-i+1} = P[X_{i,j}=1]$



$j = i + \Delta$

$P[X_{i,j}] = P[X_{i,i+\Delta}] = \frac{2}{\Delta+1}$

$E[Y] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P[X_{ij}=1] = \sum_i \sum_j \frac{2}{j-i+1}$

$\leq \sum_{i=1}^{n-1} \sum_{\Delta=1}^{n} \frac{2}{\Delta+1}$

$\leq 2n \sum_{\Delta=1}^{n} \frac{1}{\Delta} \leq 2n H_n \leq 2n(\ln n + 1)$

∎

High probability Quick Sort

Claim Probability QS performs more than $20 n \log n$ comparisions is $\leq \frac{1}{n^2}$