ANALYSIS OF QUICKSORT

Let $\{z_1\, z_2\, z_3 \ldots z_n\}$ be the given data arranged in sorted order.
unknown

$$X_{ij} = \begin{cases} 1 & \text{if } z_i \text{ is } \underline{\text{ever}} \text{ compared to } z_j \\ 0 & \text{otherwise} \end{cases}$$

$$X = \text{total \# comparisons} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}$$

Every x_i & x_j is compared once or never.

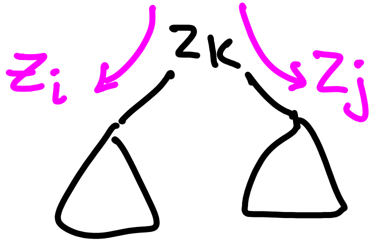
We are interested in $E[X]$

$$E[X] = E[\text{total \# comparisons}] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

linearity of expectation

$$E[X] = E[\text{total \# comparisons}] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

define $Z_{ij} = \{z_i \dots z_j\}$ (subsequence of $\{z_1 \dots z_n\}$)



z_i will be compared to z_j unless any z_k $(i<k<j)$ is a pivot before them

$z_1, \dots z_{i-1}$ & $z_{j+1} \dots z_n$ are irrelevant

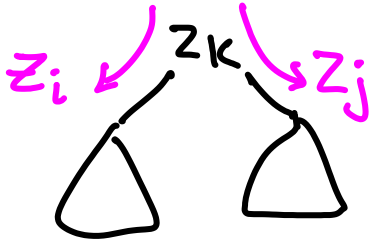
So $E[X_{ij}] = Pr\{z_i \text{ is chosen first among } Z_{ij}\}$
$+ Pr\{z_j \text{ is chosen first among } Z_{ij}\}$

$E[y] = \sum_t t \cdot P(y=t)$

For I.R.V. : $E[Y] = P[Y=1]$

$$E[X] = E[\text{total \# comparisons}] = E\left[\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}\right] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$$

define $Z_{ij} = \{z_i \ldots z_j\}$ (subsequence of $\{z_1 \ldots z_n\}$)



z_i will be compared to z_j unless any z_k $(i<k<j)$ is a pivot before them

$z_1, \ldots z_{i-1}$ & $z_{j+1} \ldots z_n$ are irrelevant

So $E[X_{ij}] = \Pr\{z_i \text{ is chosen first among } Z_{ij}\} + \Pr\{z_j \text{ is chosen first among } Z_{ij}\} = \dfrac{2}{j-i+1}$

$$E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} = \sum_{i=1}^{n-1} \sum_{t=1}^{n-i} \frac{2}{t+1} < \sum_{i=1}^{n-1} \sum_{t=1}^{n} \frac{2}{t} = \sum_{i=1}^{n-1} O(\log n) = O(n \log n)$$