Formal Analysis of Quicksort: Expected Time Complexity

define X_k : $\begin{cases} 1 & \text{if a pivot partitions array into } \underbrace{k}_{\text{left}} \text{ \& } \underbrace{n-k-1}_{\text{right}} \\ 0 & \text{otherwise} \end{cases}$

for $k = 0 \ldots n-1$

$E[X_k] = 0 \cdot P(X_k = 0) + 1 \cdot \underbrace{P(X_k = 1)}$ ← by definition

$=$ probability of generating a "k-split" $= \frac{1}{n}$ } all pivots/splits are equally likely

(= Prob. picking the k^{th} smallest)

$$T(n) = \Theta(n) + \underbrace{\left\{\begin{array}{l} T(0) + T(n-1) \\ T(1) + T(n-2) \\ T(2) + T(n-3) \\ \vdots \\ T(n-1) + T(0) \end{array}\right\}}_{}\text{n cases} \quad = \Theta(n) + \sum_{k=0}^{n-1} X_k \cdot \left(T(k) + T(n-k-1)\right)$$

$$E[T(n)] = \underbrace{E\left[\sum\right] = \sum E[(\cdot)]}_{\text{linearity of expectation}} = E[\Theta(n)] + \sum_{k=0}^{n-1} E\left[X_k \cdot \left(T(k) + T(n-k-1)\right)\right] =$$

by independence of random variables

$$= \Theta(n) + \sum E[X_k] \cdot E\left[T(k) + T(n-k-1)\right]$$

There are random variables (X_j) in recursive calls, but independent of X_k

$$\Theta(n) + \sum E[X_k] \cdot E\left[T(k) + T(n-k-1)\right]$$

$$= \Theta(n) + \sum_{k=0}^{n-1} \frac{1}{n} \cdot E\left[T(k) + T(n-k-1)\right]$$

$$= \underbrace{\frac{1}{n} \cdot \sum E\left[T(k)\right]}_{\text{symmetric: equal}} + \underbrace{\frac{1}{n} \sum E\left[T(n-k-1)\right]}$$

$$= \Theta(n) + \frac{2}{n} \sum_{k=0}^{n-1} E\left[T(k)\right]$$

$$= \underbrace{\Theta(n) + \frac{2}{n}\underbrace{\left(E\left[T(0)\right] + E\left[T(1)\right]\right)}_{\text{to make math easier later}}}_{\Theta(n)} + \frac{2}{n} \sum_{k=2}^{n-1} E\left[T(k)\right]$$

guess $E[T(n)] \leq a \cdot n \log n$

assume true for $k < n$

$$E\left[T(n)\right] = \Theta(n) + \frac{2}{n} \sum_{k=2}^{n-1} E\left[T(k)\right]$$

$$\leq \Theta(n) + \frac{2}{n} \sum_{k=2}^{n-1} a \cdot k \cdot \log k$$

Use: $\sum_{k=2}^{n-1} k \log k \leq \frac{1}{2} n^2 \log n - \frac{1}{8} n^2$ [see CLRS]

$$\leq \Theta(n) + \frac{2a}{n} \cdot \left(\frac{1}{2} n^2 \log n - \frac{1}{8} n^2\right)$$

$$= a n \log n - \frac{an}{4} + \Theta(n)$$

$$= a n \log n - \underbrace{\left(\frac{an}{4} - c \cdot n\right)}_{\text{positive if } a > 4c}$$