Quicksort

```
QUICKSORT(A, p, r)

if p < r
   then q ← PARTITION(A, p, r)
        QUICKSORT(A, p, q − 1)
        QUICKSORT(A, q + 1, r)
```

```
PARTITION(A, p, r)

y ← RANDOM(p, r)
Exchange A[y] and A[r]
x ← A[r]
i ← p − 1
for j ← p to r − 1
     do if A[j] ≤ x
           then i ← i + 1
                exchange A[i] ↔ A[j]
exchange A[i + 1] ↔ A[r]
return i + 1
```

Partition Loop Invariant

```
PARTITION(A, p, r)

y ← RANDOM(p, r)
Exchange A[y] and A[r]
x ← A[r]
i ← p − 1
for j ← p to r − 1
        do if A[j] ≤ x
                then i ← i + 1
                        exchange A[i] ↔ A[j]
exchange A[i + 1] ↔ A[r]
return i + 1
```

Loop Invariant **At the beginning of each iteration of the for loop in Partition**

1. $A[p \dots i] \le x$
2. $A[i+1 \dots j-1] > x$
3. $A[j \dots r-1]$ **is unexamined**
4. $A[r] = x$

Quicksort Analysis

- $T(n)$ is the expected running time of quicksort
- Partition takes $O(n)$ time.
- If partition is x th smallest, then

$$T(n) = T(x) + T(n-x) + O(n)$$

Quicksort Analysis

- $T(n)$ is the expected running time of quicksort
- Partition takes $O(n)$ time.
- If partition is x th smallest, then

$$T(n) = T(x) + T(n - x) + O(n)$$

For intution consider cases:

- $x = n/2$
- $x = n/10$
- $x = 1$

Cases

$$T(n) = T(x) + T(n-x) + O(n)$$

1: $x = n/2$

$$\begin{aligned} T(n) &= T(n/2) + T(n/2) + O(n) \\ &= 2T(n/2) + O(n) \\ &= O(n \log n) \end{aligned}$$

2: $x = n/10$

$$\begin{aligned} T(n) &= T(n/10) + T(9n/10) + O(n) \\ &= O(n \log n) \end{aligned}$$

3: $x = 1$

$$\begin{aligned} T(n) &= T(1) + T(n-1) + O(n) \\ &= T(n-1) + O(n) \\ &= O(n^2) \end{aligned}$$

What might this make us guess the answer is?

Following the Selection Analysis

$$T(n) = \sum_{i=1}^{n} \frac{1}{n} \left(T(i) + T(n-i) + O(n) \right)$$

could continue as in Selection.

Alternative Analysis

- We will count comparisons of data elements.
- Claim 1: The running time is dominated by comparison of data items.
- Claim 2: All comparisons are in line 6 of PARTITION, and compare some item $A[j]$ to the pivot element.
- Claim 3: Once an element is chosen as a pivot, it is never compared to any other element again.
- Claim 4: Each pair of elements is compared to each other at most once.

Analysis:

- Let the data be renamed $Z_1, \ldots, Z_n$ in sorted order.
- Use Z_{ij} to denote $Z_i, Z_{i+1}, \ldots, Z_j$
- Let X_{ij} be the indicator random variable for the comparison of Z_i to Z_j .
- Let X be the random variable counting the number of comparisons.
- By claim 4, we have

$$X = \sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij}$$

Analysis

$$X = \sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij}$$

Taking expecations

$$\begin{aligned} E[X] &= E[\sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij}] \\ &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} E[X_{ij}] \quad \textbf{linearity of expecation} \\ &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} Pr(Z_i \textbf{ is compared to } Z_j) \end{aligned}$$

What is the probability that Z_i is compared to Z_j

When is Z_i compared to Z_j ?

- **When either Z_i or Z_j is chosen as a pivot, and the other one is still in the same recursive problem.**

What is the probability that Z_i is compared to Z_j

When is Z_i compared to Z_j ?

- **When either Z_i or Z_j is chosen as a pivot, and the other one is still in the same recursive problem.**
- **Equivalently, when either Z_i or Z_j is the first element from Z_{ij} to be chosen as a pivot.**
- **What is the probability that Z_i or Z_j is the first element from Z_{ij} to be chosen as a pivot?**

What is the probability that Z_i is compared to Z_j

When is Z_i compared to Z_j ?

- **When either Z_i or Z_j is chosen as a pivot, and the other one is still in the same recursive problem.**
- **Equivalently, when either Z_i or Z_j is the first element from Z_{ij} to be chosen as a pivot.**
- **What is the probability that Z_i or Z_j is the first element from Z_{ij} to be chosen as a pivot?**
  - Z_{ij} **has** $j - i + 1$ **elements.**
  - **pivots are always chosen uniformly at random**
  - $Pr(Z_i$ **is compared to** $Z_j) = 2/(j - i + 1)$

Finishing analysis

$$
\begin{aligned}
E[X] &= E[\sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij}] \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} E[X_{ij}] \quad \textbf{linearity of expecation} \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} Pr(Z_i \textbf{ is compared to } Z_j) \\
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{2}{j-i+1}
\end{aligned}
$$

make the transformation of variables $k = j - 1 + 1$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{2}{j-i+1} \\
&= \sum_{i=1}^{n} \sum_{k=2}^{n-i+1} \frac{2}{k} \\
&\leq 2 \sum_{i=1}^{n} \ln(n-i+1) \\
&\leq 2 \sum_{i=1}^{n} \ln(n) \\
&= O(n \log n)
\end{aligned}
$$