Quicksort

$\textsc{Quicksort}(A, p, r)$
 if $p < r$
  then $q \leftarrow \textsc{Partition}(A, p, r)$
   $\textsc{Quicksort}(A, p, q-1)$
   $\textsc{Quicksort}(A, q+1, r)$

$\textsc{Partition}(A, p, r)$
 $x \leftarrow A[r]$
 $i \leftarrow p - 1$
 for $j \leftarrow p$ **to** $r - 1$
  do if $A[j] \leq x$
   then $i \leftarrow i + 1$
    exchange $A[i] \leftrightarrow A[j]$
 exchange $A[i+1] \leftrightarrow A[r]$
 return $i + 1$

Performance of QUICKSORT

Let $T(n)$ = number of $A[j] \le x$ comparisons.

Worst-case: PARTITION always returns p or r. $T(n) = T(n-1) + T(0) + n - 1$ implies $T(n)$ is $\Theta(n^2)$.

Proof: $T(0) = 0$, so recurrence simplifies to:

$$T(n) = T(n-1) + n - 1 = \sum_{k=1}^{n} (k-1)$$

Here is an upper bound on $T(n)$.

$$T(n) = \sum_{k=1}^{n} (k-1) \le n\,(n-1) < n^2$$

Here is a lower bound on $T(n)$.

$$\begin{aligned} T(n) &= \sum_{k=1}^{n} (k-1) \ge \sum_{k=\lfloor n/2 \rfloor}^{n} (k-1) \\ &\ge (n/2)\,(n/2-1) = n^2/4 - n/2 \end{aligned}$$

Best-case: PARTITION returns $\lfloor (p+r)/2 \rfloor$.

By Master Method, $T(n) = 2\,T(n/2) + n - 1$ implies $T(n) \in \Theta(n \lg n)$.

To understand average-case, consider average partition size.

Average partition: Let s_1 and s_2 = region sizes Assume that s_1 is equally likely to be any real number from 0 to n, and that $s_1 + s_2 = n$

With prob. $1/4$, $s_1 \le n/4$, and $s_2 \ge 3n/4$.
With prob. $1/4$, $n/4 \le s_1 \le n/2 \le s_2 \le 3n/4$.
With prob. $1/4$, $3n/4 \ge s_1 \ge n/2 \ge s_2 \ge n/4$.
With prob. $1/4$, $s_1 \ge 3n/4$, and $s_2 \le n/4$.

So with prob. $1/2$, $\max(s_1, s_2) \le 3n/4$.
and with prob. $1/2$, $\max(s_1, s_2) \ge 3n/4$.

This implies a median value of about $3n/4$ for the maximum of s_1 and s_2. When does median = average?

Average-Case: Each partition equally likely.

This can be guaranteed by adding the following randomization to PARTITION:

RANDOMIZED-PARTITION(A, p, r)
  $i \leftarrow$ RANDOM(p, r)
  exchange $A[r] \leftrightarrow A[i]$

  $\triangleright$ Rest of procedure same as PARTITION

This guarantees that every element is equally likely to be the pivot.

$$\begin{aligned} T(n) &= n - 1 + \frac{1}{n} \sum_{k=0}^{n-1} (T(k) + T(n-1-k)) \\ &= n - 1 + \frac{2}{n} \sum_{k=0}^{n-1} T(k) \end{aligned}$$

$$T(n-1) = n - 2 + \frac{2}{n-1} \sum_{k=0}^{n-2} T(k)$$

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

$$nT(n) = 2(n-1) + (n+1)T(n-1)$$

$$\begin{aligned}
\frac{T(n)}{n+1} &< \frac{2}{n} + \frac{T(n-1)}{n} \\
&< \frac{2}{n} + \frac{2}{n-1} + \frac{T(n-2)}{n-1} \\
&< 2 \sum_{k=2}^{n} \frac{1}{k} < 2 \ln n
\end{aligned}$$