Average-case Analysis for Insertion Sort

Let us assume that all the numbers we are sorting are distinct and that any input order is equally likely.

The sample space for the insertion sort algorithm (**IS**) is the set of possible permutations of the elements of $\{1, 2, \ldots, n\}$. Each permutation is equally likely to occur as input to **IS**. For each input π, the number of *inversions* in it is the number of pairs (i, j), where $1 \le i < j \le n$, such that j appears before i in π. Observe that the number of comparisions performed by **IS** on π is precisely the number of inversions in π.

For all $i < j$, define the 0-1 random variable X_{ij} to be 1 if and only if j appears before i in the input. Let X be the random variable denoting the total number of comparisions performed, which equals

$$\sum_{(i,j):1\le i<j\le n} X_{ij}.$$

By linearity of expectation,

$$E[X] = \sum_{(i,j):1\le i<j\le n} E[X_{ij}].$$

For any fixed $i < j$, the number of permutations in which i appears before j is equal to the number in which i appears after j. Therefore $\Pr(X_{ij} = 1) = \Pr(X_{ij} = 0) = 1/2$, implying that

$$E[X_{ij}] = 1 \cdot \Pr(X_{ij} = 1) + 0 \cdot \Pr(X_{ij} = 0) = 1/2.$$

The number of choices of (i, j) such that $1 \le i < j \le n$ is precisely $\binom{n}{2}$. Thus,

$$E[X] = \sum_{(i,j):1\le i<j\le n} E[X_{ij}] = \frac{1}{2}\binom{n}{2},$$

which is half the number of comparisions that selection sort performs on the average (in fact on every input).