Homework 1 Report

Part a) Calculating Expected value of number of duplications for an array of size n:

X is our random variable. We want to calculate the expected value of duplicate pair occurrence, $E[X]$. We will benefit from indicator random variables since it is easier to compute the result with them. Let's denote the indicator random variables with X_i. Then:

$$E[X] = E[\sum X_i]$$

Using Linearity property of Expected value:

$$E[X] = \sum E[X_i] = E[X_1] + E[X_2] + \cdots + E[X_{C(n,2)}]$$

Where n is the array size. We can easily calculate this because we know that $E[X_i] = Pr\{"i^{th}\ pair\ is\ duplicate"\}$ by the definition of indicator random variables. And $Pr\{"i^{th}\ pair\ is\ duplicate"\} = 1/n$, since each number is an integer in the interval $[1, n]$. We have $C(n, 2) = \frac{n(n-1)}{2}$ indicator random variables and all of them have the same value.

As a result, we get:

$$E[X] = \frac{1}{n} \times \frac{n(n-1)}{2} = \frac{n-1}{2}$$

Part b) Runtime complexity of the program:

In the implemented program, the outer loop (i) goes from 1 to n-1, inner loop (j) goes from i to n and inner operations take c steps. Total number of steps can be calculated as:

$$f(n) = (n-1) * c + (n-2) * c + \cdots + 1 * c = c * \frac{(n-1)(n)}{2}$$

Therefore;

$$f(n) \in O(n^2)$$