1. Linearity of expectation continued

**Example 1.** Consider a complete graph with 5 vertices (i.e., every two vertices are connected by an edge). For each of 10 edges in the graph, we toss a coin, and if heads occurs, then we erase the edge. By $X$ we denote a random variable that is equal to the number of triangles in the resulting graph.

We label the vertices with numbers 1 to 5. For every triple of vertices $(i,j,k)$ we introduce a new random variable $X_{i,j,k}$ such that

$$X_{i,j,k} = \begin{cases} 
1, & \text{if these vertices form a triangle,} \\
0, & \text{otherwise.}
\end{cases}$$

Then we can write

$$X = \sum_{1 \leq i < j < k \leq 5} X_{i,j,k}.$$ 

Using the linearity of the expectation, we obtain

$$E(X) = \sum_{1 \leq i < j < k \leq 5} E(X_{i,j,k}).$$

Consider some $i,j,$ and $k$. In order $X_{i,j,k}$ to be equal to one, all edges $i-j$, $j-k$, and $i-k$ have not to be erased. Since there are seven more edges left, this happens in $2^7$ elementary events. Then the probability of the existence of the triangle $(i,j,k)$ is equal to $2^{-7} = \frac{1}{8}$. Then $E(X_{i,j,k}) = 1 \cdot \frac{1}{8} + 0 \cdot \frac{7}{8}$. There are 10 triples $1 \leq i < j < k \leq 5$ in total, so $E(X) = \frac{10}{8} = 1.25$.

2. Simple randomized algorithm

In this section, we will design and analyse an efficient randomized algorithm for the following problem.

**Input:** an array $A$ of the length $2n$ such that exactly $n$ elements are equal to $X$ and exactly $n$ elements are equal to $Y$.

**Output:** An index $i$ (at least one) such that $A[i] = X$.

There is a straightforward deterministic algorithm solving this problem

```plaintext
function FindXDetermenistic(A[1..2n])
for i from 1 to n + 1 do
    if A[i] == X then
        return i
    end if
end for
end function
```

Its worst-case complexity (i.e., the number of operations in the worst case) is $\Theta(n)$.

Consider the following randomized algorithm

```plaintext
function FindX(A[1..2n])
for j from 1 to n do
    i = random(1, 2n)
    if A[i] == X then
        return i
    end if
end for
return FindXDetermenistic(A)
end function
```

1
This algorithm first tries to find a desired index “by chance” and if it does not succeed, it calls a deterministic algorithm. In the worst case, the number of operations will be again $\Theta(n)$.

**Statement 1.** For every array $A$ of length $2n$ with $n$ elements equal to $X$ and $n$ elements equal to $Y$, the expectation (i.e. the average) of the number of operations is $\Theta(1)$.

**Proof.** There exist constants $c_1$ and $c_2$ such that the number of operations performed by the call of FindXDeterministic$(A)$ is at most $c_1n$ and the number of operations performed by every iteration of the first loop in Algorithm 1 is at most $c_2$.

For every $1 \leq k \leq n$, we introduce an event

$$B_k = \text{“a desired index was found on } k\text{-th iteration of the first loop”}.$$

We also introduce event $C = \text{“a desired index was not found by the first loop.”}$

- The probability of $B_1$ being equal to $\frac{1}{2}$, because exactly half of indices from 1 to $2n$ would give us a result.
- If the result was not found at the first iteration (this happens in the half of cases), it will be found in the half of cases at the next iteration, so $\Pr(B_2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.
- There are only $\frac{1}{2}$ of the cases, when the answer is not found by neither of the first two iterations. In the half of these cases, a result will be found by the third iteration, so $\Pr(B_3) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{6}$.
- In general, the probability of the fact that a result will not be found by the first $k$ iterations is $\frac{1}{2^{k-1}}$. In the half of the cases, the $k$-th iteration will find a result, so $\Pr(B_k) = \frac{1}{2^{k-1}} \cdot \frac{1}{2} = \frac{1}{2^k}$.
- Analogously, $\Pr(C) = \frac{1}{2^k}$

If event $B_k$ happens, then the number of operations is $c_2k$, if $C$ happens, then the number of operations is at most $c_1n + c_2n$. Then the expectation of the number of operations is equal to

$$c_2 \Pr(B_1) + 2c_2 \Pr(B_2) + \ldots + nc_2 \Pr(B_n) + (c_1n + c_2n) \Pr(C) = c_2 \left( \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \ldots + \frac{n}{2^n} \right) + \frac{c_1n + c_2n}{2^n}.

The first summand can be found as follows

$$\frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \ldots + \frac{n}{2^n} = \frac{1}{2} + \left( \frac{1}{2^2} + \frac{1}{2^2} \right) + \left( \frac{1}{2^3} + \frac{1}{2^3} + \frac{1}{2^3} \right) + \ldots + \left( \frac{1}{2^n} + \ldots + \frac{1}{2^n} \right) = \left( \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots + \frac{1}{2^n} \right) + \ldots + \left( \frac{1}{2n-1} + \frac{1}{2n} + \frac{1}{2n} \right) + \frac{1}{2n} = \left( 1 - \frac{1}{2n} \right) + \left( \frac{1}{2} - \frac{1}{2n} \right) + \ldots + \left( \frac{1}{2n} - \frac{1}{2n} \right) = \left( 1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2n-1} \right) - \frac{n}{2^n} = 2 - \frac{n + 2}{2^n}.

Hence the expectation can be written as

$$c_1 \left( 2 - \frac{n + 2}{2^n} \right) + \frac{(c_1 + c_2)n}{2^n} = \Theta(1).$$

\[\square\]

### 3. Conditional probabilities and independence

Assume that there are two events $A$ and $B$, and we know that $B$ already happened. We would like to define what is the probability of $A$ given that $B$ happened. Since $B$ happened, the only relevant elementary events are elementary events in $B$. So it is reasonable to set the probability of $A$ given $B$ to be equal

$$\Pr(A|B) := \frac{\Pr(A \cap B)}{\Pr(B)}.$$

The above formula can be understood as “the ratio of $B$ occupied by $A$.”
**Example 2.** We roll a die twice. Let $A$ be “the number after the first roll is three” and let $B$ be “the sum of the results is seven.” Then we have

$$B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\} \text{ and } A \cap B = \{(3, 4)\}.$$ 

Hence

$$\Pr(A|B) = \frac{1}{36} \cdot \frac{6}{36} = \frac{1}{6}.$$ 

If the fact the $B$ happened does not affect the probability of $A$, then it makes sense to call $A$ and $B$ independent. This can be written as

$$\Pr(A|B) = \Pr(A) \Rightarrow \Pr(A) \cdot \Pr(B) = \Pr(A \cap B).$$ 

The formal definition is that events $A$ and $B$ are *independent* if and only if

$$\Pr(A) \cdot \Pr(B) = \Pr(A \cap B).$$ 

**Example 3.** Consider $A$ and $B$ from Example 2. Since $\Pr(A) = \frac{1}{6}$ and $\Pr(A|B) = \frac{1}{6}$, $A$ and $B$ are independent.

But if we introduce $C$ to be “the sum of the results is equal to six”, then

$$C = \{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\} \text{ and } A \cap C = \{(3, 3)\}.$$ 

Thus, $\Pr(A|C) = \frac{1}{3}$, so $A$ and $C$ are not independent.