

# Mathematical Foundations of Computer Science

## Lecture Outline

October 29, 2020

---

### Expectation

The PMF of a random variable,  $X$ , provides us with many numbers, the probabilities of all possible values of  $X$ . It would be desirable to summarize this distribution into a representative number that is also easy to compute. This is accomplished by the *expectation* of a random variable which is the weighted average (proportional to the probabilities) of the possible values of  $X$ .

**Definition.** The *expectation* of a discrete random variable  $X$ , denoted by  $\mathbf{E}[X]$ , is given by

$$\mathbf{E}[X] = \sum_i i p_X(i) = \sum_i i \Pr[X = i]$$

Intuitively,  $\mathbf{E}[X]$  is the value we would expect to obtain if we repeated a random experiment several times and took the average of the outcomes of  $X$ .

In our running example, in expectation the number of heads is given by

$$\mathbf{E}[X] = 0 \times \frac{1}{8} + 3 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} = \frac{3}{2}$$

As seen from the example, the expectation of a random variable may not be a valid value of the random variable.

**Example.** When we roll a die what is the result in expectation?

**Solution.** Let  $X$  be the random variable that denotes the result of a single roll of dice. The PMF for  $X$  is given by

$$p_X(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6.$$

The expectation of  $X$  is given by

$$\mathbf{E}[X] = \sum_{x=1}^6 p_x(x) \cdot x = \frac{1}{6} (1 + 2 + 3 + 4 + 5 + 6) = 3.5$$

**Example.** When we roll two dice what is the expected value of the sum?

**Solution.** Let  $S$  be the random variable denoting the sum. The PMF for  $S$  is given by

$$p_S(x) = \begin{cases} \frac{1}{36}, & x = 2, 12 \\ \frac{2}{36}, & x = 3, 11 \\ \frac{3}{36}, & x = 4, 10 \\ \frac{4}{36}, & x = 5, 9 \\ \frac{5}{36}, & x = 6, 8 \\ \frac{6}{36}, & x = 7 \end{cases}$$

The expectation of  $S$  is given by

$$\begin{aligned} \mathbf{E}[S] &= \sum_{x=2}^{12} p_S(x) \cdot x \\ &= \frac{1}{36} \times 2 + \frac{2}{36} \times 3 + \frac{3}{36} \times 4 + \frac{4}{36} \times 4 + \frac{5}{36} \times 6 + \frac{6}{36} \times 7 + \\ &\quad \frac{5}{36} \times 8 + \frac{4}{36} \times 9 + \frac{3}{36} \times 10 + \frac{2}{36} \times 11 + \frac{1}{36} \times 12 \\ &= \frac{252}{36} = 7 \end{aligned}$$

## Linearity of Expectation

One of the most important properties of expectation that simplifies its computation is the *linearity of expectation*. By this property, the expectation of the sum of random variables equals the sum of their expectations. This is given formally in the following theorem. I didn't cover the proof in the class but I am including it here for anyone who is interested.

**Theorem.** For any finite collection of random variables  $X_1, X_2, \dots, X_n$ ,

$$\mathbf{E} \left[ \sum_{i=1}^n X_i \right] = \sum_{i=1}^n \mathbf{E}[X_i]$$

**Proof.** We will prove the statement for two random variables  $X$  and  $Y$ . The general claim can be proven using induction.

$$\begin{aligned}
\mathbf{E}[X + Y] &= \sum_i \sum_j (i + j) \Pr[X = i \cap Y = j] \\
&= \sum_i \sum_j (i \Pr[X = i \cap Y = j] + j \Pr[X = i \cap Y = j]) \\
&= \sum_i \sum_j i \Pr[X = i \cap Y = j] + \sum_i \sum_j j \Pr[X = i \cap Y = j] \\
&= \sum_i i \sum_j \Pr[X = i \cap Y = j] + \sum_j j \sum_i \Pr[X = i \cap Y = j] \\
&= \sum_i i \Pr[X = i] + \sum_j j \Pr[Y = j] \\
&= \mathbf{E}[X] + \mathbf{E}[Y]
\end{aligned}$$

It is important to note that no assumptions have been made about the random variables while proving the above theorem. For example, the random variables do not have to be independent for linearity of expectation to be true.

**Lemma.** For any constant  $c$  and discrete random variable  $X$ ,

$$\mathbf{E}[cX] = c\mathbf{E}[X]$$

**Proof.** The lemma clearly holds for  $c = 0$ . For  $c \neq 0$

$$\begin{aligned}
\mathbf{E}[cX] &= \sum_j j \Pr[cX = j] \\
&= c \sum_j (j/c) \Pr[X = j/c] \\
&= c \sum_k k \Pr[X = k] \\
&= c\mathbf{E}[X]
\end{aligned}$$

**Example.** Using linearity of expectation calculate the expected value of the sum of the numbers obtained when two dice are rolled.

**Solution.** Let  $X_1$  and  $X_2$  denote the random variables that denote the result when die 1 and die 2 are rolled respectively. We want to calculate  $\mathbf{E}[X_1 + X_2]$ . By linearity of expectation

$$\begin{aligned}
\mathbf{E}[X_1 + X_2] &= \mathbf{E}[X_1] + \mathbf{E}[X_2] \\
&= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) + \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) \\
&= 3.5 + 3.5 \\
&= 7
\end{aligned}$$

**Example.** Suppose that  $n$  people leave their hats at the hat check. If the hats are randomly returned what is the expected number of people that get their own hat back?

**Solution.** Let  $X$  be the random variable that denotes the number of people who get their own hat back. Let  $X_i, 1 \leq i \leq n$ , be the random variable that is 1 if the  $i$ th person gets his/her own hat back and 0 otherwise. Clearly,

$$X = X_1 + X_2 + X_3 + \dots + X_n$$

By linearity of expectation we get

$$\mathbf{E}[X] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n \frac{(n-1)!}{n!} = n \times \frac{1}{n} = 1$$

**Example.** Suppose we throw  $n$  balls into  $n$  bins with the probability of a ball landing in each of the  $n$  bins being equal. What is the expected number of empty bins?

**Solution.** First Approach: The following approach was discussed in class. Let  $X$  be the random variable denoting the number of empty bins. For  $0 \leq i \leq n$ , let  $X_i$  be a random variable that is  $i$  if exactly  $i$  bins are empty and 0, otherwise. We have

$$X = \sum_{i=1}^n X_i$$

By the linearity of expectation, we have

$$\mathbf{E}[X] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n \mathbf{E}[X_i] = \sum_{i=1}^n i \Pr[X_i = i] = \sum_{i=1}^n i \Pr[X = i]$$

The last equality follows because exactly one of the  $X_i$ s will be non-zero and if  $X_i \neq 0$  then  $X = X_i$ . Note that we have not made any progress as we are back to using the original definition of expectation to solve the problem.

Second Approach: Let  $X$  be the random variable denoting the number of empty bins. Let  $X_i$  be a random variable that is 1 if the  $i$ th bin is empty and is 0 otherwise. Clearly

$$X = \sum_{i=1}^n X_i$$

By linearity of expectation, we have

$$\begin{aligned}
 \mathbf{E}[X] &= \sum_{i=1}^n \mathbf{E}[X_i] \\
 &= \sum_{i=1}^n \Pr[X_i = 1] \\
 &= \sum_{i=1}^n \left(\frac{n-1}{n}\right)^n \\
 &= \sum_{i=1}^n \left(1 - \frac{1}{n}\right)^n
 \end{aligned}$$

As  $n \rightarrow \infty$ ,  $(1 - \frac{1}{n})^n \rightarrow \frac{1}{e}$ . Hence, for large enough values of  $n$  we have

$$\mathbf{E}[X] = \frac{n}{e}$$

**Example.** The following pseudo-code computes the minimum of  $n$  distinct numbers that are stored in an array  $A$ . What is the expected number of times that the variable  $min$  is assigned a value if the array  $A$  is a random permutation of the  $n$  elements.

```

FINDMIN( $A, n$ ):
 $min \leftarrow A[1]$ 
for  $i \leftarrow 2$  to  $n$  do
    if ( $A[i] < min$ ) then
         $min \leftarrow A[i]$ 
return  $min$ 

```

**Solution.** Let  $X$  be the random variable denoting the number of times that  $min$  is assigned a value. We want to calculate  $\mathbf{E}[X]$ . Let  $X_i$  be the random variable that is 1 if  $min$  is assigned  $A[i]$  and 0 otherwise. Clearly,

$$X = X_1 + X_2 + X_3 + \cdots + X_n$$

Using the linearity of expectation we get

$$\begin{aligned}
 \mathbf{E}[X] &= \sum_{i=1}^n \mathbf{E}[X_i] \\
 &= \sum_{i=1}^n \Pr[X_i = 1]
 \end{aligned} \tag{1}$$

Note that  $\Pr[X_i = 1]$  is the probability that  $A[i]$  contains the smallest element among the elements  $A[1], A[2], \dots, A[i]$ . Since the smallest of these elements is equally likely to be in any of the first  $i$  locations, we have  $\Pr[X_i = 1] = \frac{1}{i}$ . Thus equation (1) becomes

$$\mathbf{E}[X] = \sum_{i=1}^n \frac{1}{i} = H(n) \approx \ln n + c$$

where  $c$  is a constant less than 1.

**Example.** Suppose there are  $k$  people in a room and  $n$  days in a year. On average how many pairs of people share the same birthday?

**Solution.** Let  $X$  be the random variable denoting the number of pairs of people sharing the same birthday. For any two people  $i$  and  $j$ , let  $X_{ij}$  be an indicator random variable that is 1 if  $i$  and  $j$  have the same birthday and is 0 otherwise. Clearly  $X = \sum_{i,j} X_{ij}$ . Using the linearity of expectation we get

$$\begin{aligned}\mathbf{E}[X] &= \sum_{i,j} \mathbf{E}[X_{ij}] \\ &= \sum_{i,j} \Pr[X_{ij} = 1] \\ &= \sum_{i,j} \frac{1}{n} \\ &= \frac{\binom{k}{2}}{n} \\ &= \frac{k(k-1)}{2n}\end{aligned}$$

Assuming  $n = 365$ , the smallest value of  $k$  for which the RHS is at least 1 is 28.

**Example (Markov's Inequality).** Let  $X$  be a non-negative random variable. Then for all  $a > 0$ , prove that

$$\Pr[X \geq a] \leq \frac{\mathbf{E}[X]}{a}$$