Mathematical Foundations of Computer Science
Lecture Outline
October 17, 2024

Example. An urn contains 5 white and 10 black balls. A fair die is rolled and that number of balls are chosen from the urn.
(a) What is the probability that all of the balls selected are white?
(b) What is the conditional probability that the die landed on 3 if all the balls selected are white?

Solution. We will consider the following events.

W: event that all of the balls chosen are white.
D_i: event that the die landed on i, $1 \leq i \leq 6$.

(a) We want to find $\Pr[W]$. We can do this as follows.

$$\begin{aligned}
\Pr[W] &= \sum_{i=1}^{6} \Pr[W \cap D_i] \\
&= \sum_{i=1}^{6} \Pr[D_i] \Pr[W|D_i] \\
&= \sum_{i=1}^{6} \frac{1}{6} \frac{\binom{5}{i}}{\binom{15}{i}} \\
&= \frac{1}{6}\left(\frac{5}{15} + \frac{10}{105} + \frac{10}{455} + \frac{5}{1365} + \frac{1}{3003}\right) \\
&= 0.075
\end{aligned}$$

(b) We want to find $\Pr[D_3|W]$. This can be done as follows.

$$\begin{aligned}
\Pr[D_3|W] &= \frac{\Pr[D_3 \cap W]}{\Pr[W]} \\
&= \frac{\Pr[D_3] \times \Pr[W|D_3]}{\Pr[W]} \\
&= \frac{1/6 \times \binom{5}{3}/\binom{15}{3}}{0.075} \\
&= \frac{1/6 \times 10/455}{0.075} \\
&= \frac{0.00366}{0.075} \\
&= 0.048
\end{aligned}$$

Example. Consider the experiment in which we roll a dice twice. Consider the following events.

> A: event that the first roll results in a 1, 2, or a 3.
> B: event that the first roll results in a 3, 4, or a 5.
> C: event that the sum of the two rolls is a 9

Are events A, B, and C mutually independent?

Solution. We show below that the events are not mutually independent as they are not pairwise independent.

$$\begin{aligned}
A &= \{(i,j)\,|\,1 \le i \le 3 \text{ and } 1 \le j \le 6\} \\
B &= \{(i,j)\,|\,3 \le i \le 5 \text{ and } 1 \le j \le 6\} \\
C &= \{(3,6),(6,3),(4,5),(5,4)\} \\
A \cap B &= \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\} \\
A \cap C &= \{(3,6)\} \\
B \cap C &= \{(3,6),(4,5),(5,4)\} \\
A \cap B \cap C &= \{(3,6)\} \\
\Pr[A] &= 1/2 \\
\Pr[B] &= 1/2 \\
\Pr[C] &= 1/9 \\
\Pr[A \cap B \cap C] &= 1/36 = \Pr[A] \cdot \Pr[B] \cdot \Pr[C] \\
\Pr[A \cap B] &= 1/6 \ne \Pr[A] \cdot \Pr[B] \\
\Pr[A \cap C] &= 1/36 \ne \Pr[A] \cdot \Pr[C] \\
\Pr[B \cap C] &= 3/36 \ne \Pr[B] \cdot \Pr[C]
\end{aligned}$$

Random Variables

In an experiment we are often interested in some value associated with an outcome as opposed to the actual outcome itself. For example, consider an experiment that involves tossing a coin three times. We may not be interested in the actual head-tail sequence that results but be more interested in the number of heads that occur. These quantities of interest are called *random variables.*

Definition. A *random variable* X on a sample space Ω is a real-valued function that assigns to each sample point $\omega \in \Omega$ a real number $X(\omega)$.

In this course we will study discrete random variables which are random variables that take on only a finite or countably infinite number of values.

For a discrete random variable X and a real value a, the event "X=a" is the set of outcomes in Ω for which the random variable assumes the value a, i.e., $X = a \equiv \{\omega \in \Omega | X(\omega) = a\}$. The probability of this event is denoted by

$$\Pr[X = a] = \sum_{\omega \in \Omega : X(\omega) = a} \Pr[\omega]$$

Definition. The *distribution* or the *probability mass function*(PMF) of a random variable X gives the probabilities for the different possible values of X. Thus, if x is a value that X can assume then $p_X(x)$ is the probability mass of X and is given by

$$p_X(x) = \Pr[X = x]$$

Observe that $\sum_x p_X(x) = \sum_x \Pr[X = x] = 1$. This is because the events $X = x$ are disjoint and hence partition the sample space Ω.

Consider the experiment of tossing three fair coins. Let X be the random variable that denotes the number of heads that result. The PMF or the distribution of X is given below.

$$p_X(x) = \begin{cases} 1/8 & \text{if } x = 0 \text{ or } x = 3 \\ 3/8 & \text{otherwise} \end{cases}$$

The definition of independence that we developed for events extends to random variables.

Definition. Two random variables X and Y are independent if and only if

$$\Pr[(X = x) \cap (Y = y)] = \Pr[X = x] \times \Pr[Y = y]$$

for all values x and y. In other words, two random variables X and Y are independent if every event determined by X is independent of every event determined by Y.

Similarly, random variables $X_1, X_2, \ldots, X_k$ are mutually independent if and only if, for any subset $I \subseteq [1, k]$ and any values $x_i, i \in I$,

$$\Pr[\cap_{i \in I} X_i = x_i] = \prod_{i \in I} \Pr[X_i = x_i]$$

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 $X_1, X_2, \ldots, X_n$,

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

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 ith person gets his/her own hat back and 0 otherwise. Clearly,

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

By linearity of expectation we get

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