Example. Suppose that a fair coin is tossed twice. Let $A$ be the event that a head is obtained on the first toss, $B$ be the event that a head is obtained on the second toss, and $C$ be the event that either two heads or two tails are obtained. (a) Are events $A, B, C$ pairwise independent? (b) Are they mutually independent?

Solution. Note that $\Omega = \{HH, HT, TH, TT\}$. $A = \{HH, HT\}$, $B = \{HH, TH\}$, $C = \{HH, TT\}$, $A \cap B = \{HH\}$, $A \cap C = \{HH\}$, $B \cap C = \{HH\}$, $A \cap B \cap C = \{HH\}$. The probabilities of the relevant events are as follows.

\[
\begin{align*}
\text{Pr}[A] &= 1/2 \\
\text{Pr}[B] &= 1/2 \\
\text{Pr}[C] &= 1/2 \\
\text{Pr}[A \cap B] &= 1/4 = \text{Pr}[A] \text{Pr}[B] \\
\text{Pr}[A \cap C] &= 1/4 = \text{Pr}[A] \text{Pr}[C] \\
\text{Pr}[B \cap C] &= 1/4 = \text{Pr}[B] \text{Pr}[C] \\
\text{Pr}[A \cap B \cap C] &= 1/4 \neq \text{Pr}[A] \text{Pr}[B] \text{Pr}[C]
\end{align*}
\]

Thus we see that $A, B, C$ are pairwise independent but not mutually independent.

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.

\[ A = \{(i, j) | 1 \leq i \leq 3 \text{ and } 1 \leq j \leq 6 \} \]
\[ B = \{(i, j) | 3 \leq i \leq 5 \text{ and } 1 \leq j \leq 6 \} \]
\[ C = \{(3, 6), (6, 3), (4, 5), (5, 4)\} \]
\[ A \cap B = \{(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)\} \]
\[ A \cap C = \{(3, 6)\} \]
\[ B \cap C = \{(3, 6), (4, 5), (5, 4)\} \]
\[ A \cap B \cap C = \{(3, 6)\} \]
\[ \Pr[A] = \frac{1}{2} \]
\[ \Pr[B] = \frac{1}{2} \]
\[ \Pr[C] = \frac{1}{9} \]
\[ \Pr[A \cap B \cap C] = \frac{1}{36} = \Pr[A] \cdot \Pr[B] \cdot \Pr[C] \]
\[ \Pr[A \cap B] = \frac{1}{6} \neq \Pr[A] \cdot \Pr[B] \]
\[ \Pr[A \cap C] = \frac{1}{36} \neq \Pr[A] \cdot \Pr[C] \]
\[ \Pr[B \cap C] = \frac{3}{36} \neq \Pr[B] \cdot \Pr[C] \]

**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 \( \Omega \) 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 \( \Omega \) 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 $\Omega$.

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 $\mathbb{E}[X]$, is given by

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

Intuitively, $\mathbb{E}[X]$ 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

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

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

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

The expectation of $X$ is given by

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

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

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

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 $X_1, X_2, \ldots, X_n$,

$$E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i]$$
Proof. We will prove the statement for two random variables $X$ and $Y$. The general claim can be proven using induction.

$$E[X + Y] = \sum_i \sum_j (i + j) \Pr[X = i \cap Y = j]$$

$$= \sum_i \sum_j (i \Pr[X = i \cap Y = j] + j \Pr[X = i \cap Y = j])$$

$$= \sum_i \sum_j i \Pr[X = i \cap Y = j] + \sum_i \sum_j j \Pr[X = i \cap Y = j]$$

$$= \sum_i \Pr[X = i] + \sum_j j \Pr[Y = j]$$

$$= E[X] + E[Y]$$

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 $X$,

$$E[cX] = cE[X]$$

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

$$E[cX] = \sum_j j \Pr[cX = j]$$

$$= c \sum_j (j/c) \Pr[X = j/c]$$

$$= c \sum_k k \Pr[X = k]$$

$$= cE[X]$$

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 $E[X_1 + X_2]$. By linearity of expectation

$$E[X_1 + X_2] = E[X_1] + E[X_2]$$

$$= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) + \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6)$$

$$= 3.5 + 3.5$$

$$= 7$$
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,

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

By linearity of expectation we get

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

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

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

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 $\overline{X_i}$ 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

\[ E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \Pr[X_i = 1] = \sum_{i=1}^{n} \left(\frac{n-1}{n}\right)^n = \sum_{i=1}^{n} \left(1 - \frac{1}{n}\right)^n \]

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

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

**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 \( \text{min} \) is assigned a value if the array \( A \) is a random permutation of the \( n \) elements.

```
\text{FindMin}(A, n):
    \text{min} \leftarrow A[1]
    \text{for } i \leftarrow 2 \text{ to } n \text{ do}
        \text{if } (A[i] < \text{min}) \text{ then}
            \text{min} \leftarrow A[i]
    \text{return } \text{min}
```

**Solution.** Let \( X \) be the random variable denoting the number of times that \( \text{min} \) is assigned a value. We want to calculate \( E[X] \). Let \( X_i \) be the random variable that is 1 if \( \text{min} \) 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

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

(1)

Note that \( \Pr[X_i = 1] \) is the probability that \( A[i] \) contains the smallest element among the elements \( A[1], A[2], \ldots, A[i] \). 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

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

where \( c \) is a constant less than 1.