1. **Random Variable.**

Given a random experiment with outcomes $\Omega$, a *real valued random variable* $X$ defined over this experiment is a mapping $X : \Omega \rightarrow \mathbb{R}$. An *integer valued random variable* $X$ is a mapping from $X : \Omega \rightarrow \mathbb{Z}$.

*Examples:*

- We toss a fair coin. $X(\text{heads}) = 0$ and $X(\text{tails}) = 1$. This is a $\{0, 1\}$-random variable, or a Boolean random variable. Also called a *Bernoulli* random variable.
- We roll a fair die. $X$ takes the value on the face of the die.
- We roll two fair dice. $X$ takes the value of the sum. In this case, $X = Y + Z$ where $Y, Z$ are random variables of the kind from the previous bullet point.
- Given any event $E$, there is an associated random variable called the *indicator random variable* denoted as $1_E$, where $1_E(\omega) = 1$ if $\omega \in E$, and 0 otherwise.

2. **Events associated with random variables.**

Given a random variable $X$, we can associate many events and ask for their probabilities. For instance, we can ask $\Pr[X = x]$. More precisely, this is a shorthand for saying $\sum_{\omega \in \Omega : X(\omega) = x} \Pr[\omega]$.

Similarly, $\Pr[X \geq k]$ is a shorthand for saying $\sum_{\omega \in \Omega : X(\omega) \geq k} \Pr[\omega]$.

3. **Expectation of a Random Variable.**

*Theorem 1.* The expectation of a random variable $X$ is defined to be

$$\text{Exp}[X] = \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega] = \sum_{x \in \text{range}(X)} x \cdot \Pr[X = x]$$

*Remark:* The expectation is therefore often thought of as an inner-product (aka dot-product) of two vectors. These vectors have $|\Omega|$ dimensions. One vector is $(X(\omega) : \omega \in \Omega)$, and the other is $(\Pr[\omega] : \omega \in \Omega)$. This dot-product view is often useful (although, sadly, we may not see its ramifications in this course).

*Examples:*
• We toss a fair coin. \(X(\text{heads}) = 0\) and \(X(\text{tails}) = 1\). This is a \(\{0, 1\}\)-random variable, or a Boolean random variable. Also called a Bernoulli random variable.

\[
\text{Exp}[X] = 0 \cdot \Pr[X = 0] + 1 \cdot \Pr[X = 1] = 1/2
\]

Indeed, if the coin were not fair, and the probability that tails would come with probability \(p\), then \(\text{Exp}[X] = p\).

• We roll a fair die. \(X\) takes the value on the face of the die.

\[
\text{Exp}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5
\]

• We roll two fair dice. \(X\) takes the value of the sum. In this case, \(X = Y + Z\) where \(Y, Z\) are random variables of the kind from the previous bullet point.

This is requires a little work. The range of \(X\) is \(\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}\). We can calculate the probabilities for each (remember, it is not uniform), and then do the calculation.

**Exercise:** Please do the calculation.

We get the answer 7. Did you?

• Given any event \(E\), there is an associated random variable called the indicator random variable denoted as \(1_E\), where \(1_E(\omega) = 1\) if \(\omega \in E\), and 0 otherwise.

\[
\text{Exp}[1_E] = 0 \cdot \Pr[\neg E] + 1 \cdot \Pr[E] = \Pr[E]
\]

This is quite important. Why? Because it turns a probability calculation (the RHS) into an expectation calculation. As we show below, calculating expectations is often easier than calculating probabilities.

**Exercise:** Suppose you have a fair coin. Construct the following random variable \(Z\) whose range is \(\mathbb{N}\). You keep tossing the fair coin till you get a heads. \(Z\) is the number of times you have tossed the coin. What is \(\text{Exp}[Z]\)?

4. **Multiplication by a scalar.** If \(X\) is a random variable, and \(c\) is a “scalar” (a constant), then \(Z = c \cdot X\) is another random variable. \(\text{Exp}[c \cdot X] = c \cdot \text{Exp}[X]\).

**Exercise:** Prove this.

5. **Linearity of Expectation.** This is one of the most powerful equations in all of probability. Literally. It states the following. It literally has a four line proof.

**Theorem 2.** For any two random variables \(X\) and \(Y\), let \(Z := X + Y\). Then,

\[
\text{Exp}[Z] = \text{Exp}[X] + \text{Exp}[Y]
\]
Proof.

\[
\begin{align*}
\mathbb{E}[Z] &= \sum_{\omega \in \Omega} Z(\omega) \Pr[\omega] \quad \text{Definition of Expectation} \\
&= \sum_{\omega \in \Omega} (X(\omega) + Y(\omega)) \Pr[\omega] \quad \text{Definition of } Z \\
&= \sum_{\omega \in \Omega} X(\omega) \Pr[\omega] + \sum_{\omega \in \Omega} Y(\omega) \cdot \Pr[\omega] \quad \text{Distributivity} \\
&= \mathbb{E}[X] + \mathbb{E}[Y] \quad \text{Definition of Expectation}
\end{align*}
\]

As a corollary, we get:

**Theorem 3.** For any \(k\) random variables \(X_1, X_2, \ldots, X_k\),

\[
\mathbb{E}\left[ \sum_{i=1}^{k} X_i \right] = \sum_{i=1}^{k} \mathbb{E}[X_i]
\]

**Examples of applications.**

(a) We roll two fair dice. \(X\) takes the value of the sum. In this case, \(X = Y + Z\) where \(Y, Z\) are random variables of the kind from the previous bullet point.

Tailor-made application. \(\mathbb{E}[Y] = \mathbb{E}[Z] = 3.5\), the expected value of a single roll of a die. Thus, \(\mathbb{E}[X] = \mathbb{E}[Y + Z] = 7\) by linearity of expectation.

(b) We have a biased coin which lands heads with probability \(p\). We toss it 100 times. Let \(X\) be the number of heads we see. What is \(\mathbb{E}[X]\)?

**Remark:** Try doing this the “first-principle” way. That is, for each \(0 \leq k \leq 100\), figure out the probability \(\Pr[X = k]\) (that is, the probability we get exactly \(k\) heads), and then sum \(\sum_{k=0}^{100} k \cdot \Pr[X = k]\). Please try it; feel the sweat needed to do this. It will make you appreciate the next three lines more!

Define new random variables; define \(X_i\) to take the value 1 if the \(i\)th toss is heads, and 0 otherwise. Note, \(X = X_1 + X_2 + \cdots + X_{100}\). Note, \(\mathbb{E}[X_i] = p\) (it is a Bernoulli random variable). Thus, linearity of expectation gives \(\mathbb{E}[X] = 100p\).

(c) \(n\) people checked in their hats, but on their way out, were handed back hats randomly. What is the expected number of people who get their correct hats?

Define \(X_i\) to be 1 if the \(i\)th person gets his or her back correctly. What is \(\mathbb{E}[X_i]\)? It is \(1/n\); it is the probability that \(\sigma(i) = i\) for a random ordering \(\sigma\). Let \(Z = \sum_{i=1}^{n} X_i\). Note, \(Z\) is the number of people who get their correct hats. By linearity of expectation, \(\mathbb{E}[Z] = 1\).
(d) In a party of \( n \) people there are some pairs of people who are friends, and some pairs who are not. In all there are \( m \) pairs of friends. The host randomly divides the party by taking each person and sending them left or right at the toss of a fair coin. How many friends are sent apart (in expectation)?

**Remark:** A graph is randomly split into two. How many edges, in expectation, have endpoints in different parts?

For each pair of friends \((u, v)\), define \( X_{uv} \) which takes the value 1 if \( u \) and \( v \) are split, and takes the value 0 if \( u \) and \( v \) are not split. The probability \( u \) and \( v \) are split is \( \frac{1}{2} \) (either \( u \) is sent left, \( v \) is sent right, or vice-versa). Thus, \( \mathbb{E}[X_{uv}] = \frac{1}{2} \). Define \( Z = \sum_{(u,v) \text{: friends}} X_{uv} \); \( Z \) is the number of friends sent apart. \( \mathbb{E}[Z] = \sum_{(u,v) \text{: friends}} \mathbb{E}[X_{uv}] = \frac{m}{2} \). In expectation, half the friendships are sundered apart.

(e) In an ordering \( \sigma \) of \( (1, 2, \ldots, n) \), an inversion is a pair \( i < j \) such that \( \sigma(i) > \sigma(j) \). How many inversions, in expectation, are there in a random permutation?

Let \( \sigma \) be a random permutation. Define the indicator random variable \( X_{ij} \) for \( i < j \), which takes the value 1 if \( \sigma(i) > \sigma(j) \), and 0 otherwise. Note that \( \Pr[X_{ij} = 1] = \frac{1}{2} \); there are equally many orderings with \( \sigma(i) > \sigma(j) \) as \( \sigma(i) < \sigma(j) \). Now note that \( Z = \sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij} \) is the number of inversions in \( \sigma \). Thus, \( \mathbb{E}[Z] = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}[X_{ij}] = \frac{1}{2} \cdot \frac{n(n-1)}{2} \).