

3.1: Introduction to Random Variables

Now that we have formally defined probability and the underlying structure, we add another layer: *random variables*. Random variables allow characterization of outcomes, so that we do not need to focus on each outcome specifically. We begin with the formal definition.

Definition 3.1.1

A **random variable** is a function from a sample space S to the real numbers \mathbb{R} . We denote random variables with capital letters, e.g.,

$$X : S \rightarrow \mathbb{R}.$$

Informally, a random variable assigns numbers to outcomes in the sample space. So, instead of focusing on the outcomes themselves, we highlight a specific characteristic of the outcomes.

Example 3.1.1

Consider again the context of [Example 1.1.1](#), where we recorded the sequence of heads and tails in two tosses of a fair coin. The sample space for this random experiment is given by

$$S = \{hh, ht, th, tt\}.$$

Suppose we are only interested in tosses that result in heads. We can define a random variable X that tracks the number of heads obtained in an outcome. So, if outcome hh is obtained, then X will equal 2. Formally, we denote this as follows:

$$\begin{aligned} X : S &\rightarrow \mathbb{R} \\ s &\mapsto \text{number of } h\text{'s in } s \end{aligned}$$

Since there are only four outcomes in S , we can list the value of X for each outcome individually:

inputs: S	function: X	outputs: \mathbb{R}
hh	\xrightarrow{X}	2
th	\xrightarrow{X}	1
ht	\xrightarrow{X}	1
tt	\xrightarrow{X}	0

We can also write the above as follows:

$$X(hh) = 2, \quad X(ht) = X(th) = 1, \quad X(tt) = 0.$$

The advantage to defining the random variable X in this context is that the two outcomes ht and th are both assigned a value of 1, meaning we are not focused on the actual sequence of heads and tails that resulted in obtaining one heads.

In [Example 3.1.1](#), note that the random variable we defined only equals one of three possible values: 0, 1, 2. This is an example of what we call a *discrete* random variable. We will also encounter another type of random variable: *continuous*. The next definitions make precise what we mean by these two types.

Definition 3.1.2

A **discrete random variable** is a random variable that has only a finite or countably infinite (think integers or whole numbers) number of possible values.

Definition 3.1.3

A **continuous random variable** is a random variable with infinitely many possible values (think an interval of real numbers, e.g., $[0, 1]$).

In this chapter, we take a closer look at discrete random variables, then in [Chapter 4](#) we consider continuous random variables.

This page titled [3.1: Introduction to Random Variables](#) is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Kristin Kuter](#).