

Random Variables

A Random Variable (RV) is a variable that probabilistically takes on a value and they are one of the most important constructs in all of probability theory. You can think of an RV as being like a variable in a programming language, and in fact random variables are just as important to probability theory as variables are to programming. Random Variables take on values, have types and have domains over which they are applicable.

Random variables work with all of the foundational theory we have built up to this point. We can define events that occur if the random variable takes on values that satisfy a numerical test (eg does the variable equal 5, is the variable less than 8).

Lets look at a first example of a random variable. Say we flip three fair coins. We can define a random variable Y to be the total number of “heads” on the three coins. We can ask about the probability of Y taking on different values using the following notation:

Let Y be the number of heads on three coin flips

$P(Y = 0) = 1/8$	(T, T, T)
$P(Y = 1) = 3/8$	(H, T, T), (T, H, T), (T, T, H)
$P(Y = 2) = 3/8$	(H, H, T), (H, T, H), (T, H, H)
$P(Y = 3) = 1/8$	(H, H, H)
$P(Y \geq 4) = 0$	

Even though we use the same notation for random variables and for events (both use capital letters) they are distinct concepts. An event is a scenario, a random variable is an object. The scenario where a random variable takes on a particular value (or range of values) is an event. When possible, I will try and use letters E,F,G for events and X,Y,Z for random variables.

Using random variables is a convenient notation technique that assists in decomposing problems. There are many different types of random variables (indicator, binary, choice, Bernoulli, etc). The two main families of random variable types are discrete and continuous. Discrete random variables can only take on integer values. Continuous random variables can take on decimal values. We are going to develop our intuitions using discrete random variable and then introduce continuous.

Properties of random variables

There are many properties of a random variable some of which we will dive into extensively. Here is a brief summary. Each random variable has:

Property	Notation Example	Description
Meaning		A semantic description of the random variable
Symbol	X	A letter used to denote the random variable
Support or Range	$\{0, 1, \dots, 3\}$	the values the random variable can take on
Distribution Function (PMF or PDF)	$P(X = x)$	A function which maps values the RV can take on to likelihood.
Expectation	$E[X]$	A weighted average

Property	Notation Example	Description
Variance	$\text{Var}(X)$	A measure of spread
Standard Deviation	$\text{Std}(X)$	The square root of variance
Mode		The most likely value of the random variable

You should set a goal of deeply understanding what each of these properties mean. There are many more properties than the ones in the table above: properties like [entropy](#), [median](#), [skew](#), [kurtosis](#).

Random variables vs Events

Random variables and events are two different concepts. An event is an outcome, or a set of outcomes, to an experiment. A random variable is a more like an experiment -- it will take on an outcome eventually. Probabilities are over events, so if you want to talk about probability in the context of a random variable, you must construct an event. You can make events by using any of the [Relational Operators](#): $<$, \leq , $>$, \geq , $=$, or \neq (not equal to). This is analogous to coding where you can use relational operators to create boolean expressions from numbers.

Lets continue our example of the random variable Y which represents the number of heads on three coin flips. Here are some events using the variable Y :

Event	Meaning	Probability Statement
$Y = 1$	Y takes on the value 1 (there was one heads)	$P(Y = 1)$
$Y < 2$	Y takes on 0 or 1 (note this Y can't be negative)	$P(Y < 2)$
$X > Y$	X takes on a value greater than the value Y takes on.	$P(X > Y)$
$Y = y$	Y takes on a value represented by non-random variable y	$P(Y = y)$

You will see many examples like this last one, $P(Y = y)$, in this text book as well as in scientific and math research papers. It allows us to talk about the likelihood of Y taking on a value, in general. For example, later in this book we will derive that for three coin flips where Y is the number of heads, the probability of getting exactly y heads is:

$$P(Y = y) = \frac{0.75}{y!(3 - y)!} \quad \text{If } 0 \leq y \leq 3$$

This statement above is a function which takes in a parameter y as input and returns the numeric probability $P(Y = y)$ as output. This particular expression allows us to talk about the probability that the number of heads is 0, 1, 2 or 3 all in one expression. You can plug in any one of those values for y to get the corresponding probability. It is customary to use lower-case symbols for non-random values. The use of an equals sign in the "event" can be confusing. For example what does this expression say $P(Y = 1) = 0.375$? It says that the probability that "Y takes on the value 1" is 0.375. For discrete random variables this function is called the "[probability mass function](#)" and it is the topic of our next chapter.