2

Probability Theory on Coin Toss Space

2.1 Finite Probability Spaces

A finite probability space is used to model a situation in which a random experiment with finitely many possible outcomes is conducted. In the context of the binomial model of the previous chapter, we tossed a coin a finite number of times. If, for example, we toss the coin three times, the set of all possible outcomes is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}. \tag{2.1.1}$$

Suppose that on each toss the probability of a head (either actual or risk-neutral) is p and the probability of a tail is $q = 1-p$. We assume the tosses are independent, and so the probabilities of the individual elements ω (sequences of three tosses $\omega = \omega_1\omega_2\omega_3$) in Ω are

$$\begin{aligned} &\mathbb{P}(HHH) = p^3,\ \mathbb{P}(HHT) = p^2q, \mathbb{P}(HTH) = p^2q, \mathbb{P}(HTT) = pq^2,\\ &\mathbb{P}(THH) = p^2q,\ \mathbb{P}(THT) = pq^2,\ \mathbb{P}(TTH) = pq^2,\ \ \mathbb{P}(TTT) = q^3. \end{aligned} \tag{2.1.2}$$

The subsets of Ω are called *events*, and these can often be described in words as well as in symbols. For example, the event

$$\begin{aligned} \text{“The first toss is a head”} &= \{\omega \in \Omega; \omega_1 = H\}\\ &= \{HHH, HHT, HTH, HTT\} \end{aligned}$$

has, as indicated, descriptions in both words and symbols. We determine the probability of an event by summing the probabilities of the elements in the event, i.e.,

$$\begin{aligned} \mathbb{P}(\text{First toss is a head}) &= \mathbb{P}(HHH) + \mathbb{P}(HHT) + \mathbb{P}(HTH) + \mathbb{P}(HTT)\\ &= (p^3 + p^2q) + (p^2q + pq^2)\\ &= p^2(p+q) + pq(p+q) \end{aligned}$$

$$
\begin{aligned}
&= p^2 + pq \\
&= p(p+q) \\
&= p. \qquad (2.1.3)
\end{aligned}
$$

Thus, the mathematics agrees with our intuition.

With mathematical models, it is easy to substitute our intuition for the mathematics, but this can lead to trouble. We should instead verify that the mathematics and our intuition agree; otherwise, either our intuition is wrong or our model is inadequate. If our intuition and the mathematics of a model do not agree, we should seek a reconciliation before proceeding. In the case of (2.1.3), we set out to build a model in which the probability of a head on each toss is p, we proposed doing this by defining the probabilities of the elements of Ω by (2.1.2), and we further defined the probability of an event (subset of Ω) to be the sum of the probabilities of the elements in the event. These definitions force us to carry out the computation (2.1.3) as we have done, and we need to do this computation in order to check that it gets the expected answer. Otherwise, we would have to rethink our mathematical model for the coin tossing.

We generalize slightly the situation just described, first by allowing Ω to be any finite set, and second by allowing some elements in Ω to have probability zero. These generalizations lead to the following definition.

Definition 2.1.1. *A* finite probability space *consists of a* sample space Ω *and a* probability measure $\mathbb{P}$. *The sample space Ω is a nonempty finite set and the probability measure $\mathbb{P}$ is a function that assigns to each element ω of Ω a number in $[0,1]$ so that*

$$\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1. \qquad (2.1.4)$$

An event *is a subset of Ω, and we define the probability of an event A to be*

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega). \qquad (2.1.5)$$

As mentioned before, this is a model for some random experiment. The set Ω is the set of all possible outcomes of the experiment, $\mathbb{P}(\omega)$ is the probability that the particular outcome ω occurs, and $\mathbb{P}(A)$ is the probability that the outcome that occurs is in the set A. If $\mathbb{P}(A) = 0$, then the outcome of the experiment is sure not to be in A; if $\mathbb{P}(A) = 1$, then the outcome is sure to be in A. Because of (2.1.4), we have the equation

$$\mathbb{P}(\Omega) = 1, \qquad (2.1.6)$$

i.e., the outcome that occurs is sure to be in the set Ω. Because $\mathbb{P}(\omega)$ can be zero for some values of ω, we are permitted to put in Ω even some outcomes of the experiment that are sure not to occur. It is clear from (2.1.5) that if A and B are disjoint subsets of Ω, then

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B). \qquad (2.1.7)$$

2.2 Random Variables, Distributions, and Expectations

A random experiment generally generates numerical data. This gives rise to the concept of a random variable.

Definition 2.2.1. *Let* $(\Omega, \mathbb{P})$ *be a finite probability space. A* random variable *is a real-valued function defined on* Ω. *(We sometimes also permit a random variable to take the values* $+\infty$ *and* $-\infty$.*)*

Example 2.2.2 (Stock prices). Recall the space Ω of three independent coin-tosses (2.1.1). As in Figure 1.2.2 of Chapter 1, let us define stock prices by the formulas

$$S_0(\omega_1\omega_2\omega_3) = 4 \text{ for all } \omega_1\omega_2\omega_3 \in \Omega_3,$$

$$S_1(\omega_1\omega_2\omega_3) = \begin{cases} 8 & \text{if } \omega_1 = H, \\ 2 & \text{if } \omega_1 = T, \end{cases}$$

$$S_2(\omega_1\omega_2\omega_3) = \begin{cases} 16 & \text{if } \omega_1 = \omega_2 = H, \\ 4 & \text{if } \omega_1 \neq \omega_2, \\ 1 & \text{if } \omega_1 = \omega_2 = T, \end{cases}$$

$$S_3(\omega_1\omega_2\omega_3) = \begin{cases} 32 & \text{if } \omega_1 = \omega_2 = \omega_3 = H, \\ 8 & \text{if there are two heads and one tail}, \\ 2 & \text{if there is one head and two tails}, \\ .50 & \text{if } \omega_1 = \omega_2 = \omega_3 = T. \end{cases}$$

Here we have written the arguments of S_0, S_1, S_2, and S_3 as $\omega_1\omega_2\omega_3$, even though some of these random variables do not depend on all the coin tosses. In particular, S_0 is actually not random because it takes the value 4, regardless of how the coin tosses turn out; such a random variable is sometimes called a *degenerate random variable*. □

It is customary to write the argument of random variables as ω, even when ω is a sequence such as $\omega = \omega_1\omega_2\omega_3$. We shall use these two notations interchangeably. It is even more common to write random variables without any arguments; we shall switch to that practice presently, writing S_3, for example, rather than $S_3(\omega_1\omega_2\omega_3)$ or $S_3(\omega)$.

According to Definition 2.2.1, a random variable is a function that maps a sample space Ω to the real numbers. The *distribution* of a random variable is a specification of the probabilities that the random variable takes various values. *A random variable is not a distribution*, and *a distribution is not a random variable.* This is an important point when we later switch between the actual probability measure, which one would estimate from historical data, and the risk-neutral probability measure. The change of measure will change distributions of random variables but not the random variables themselves. We make this distinction clear with the following example.

Example 2.2.3. Toss a coin three times, so the set of possible outcomes is

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

Define the random variables

$$X = \text{Total number of heads}, \qquad Y = \text{Total number of tails}.$$

In symbols,

$$\begin{aligned} X(HHH) &= 3, \\ X(HHT) = X(HTH) = X(THH) &= 2, \\ X(HTT) = X(THT) = X(TTH) &= 1, \\ X(TTT) &= 0, \\ \\ Y(TTT) &= 3, \\ Y(TTH) = Y(THT) = Y(HTT) &= 2, \\ Y(THH) = Y(HTH) = Y(HHT) &= 1, \\ Y(HHH) &= 0. \end{aligned}$$

We do not need to know probabilities of various outcomes in order to specify these random variables. However, once we specify a probability measure on Ω, we can determine the distributions of X and Y. For example, if we specify the probability measure $\widetilde{\mathbb{P}}$ under which the probability of head on each toss is $\frac{1}{2}$ and the probability of each element in Ω is $\frac{1}{8}$, then

$$\begin{aligned} \widetilde{\mathbb{P}}\{\omega \in \Omega; X(\omega) = 0\} &= \widetilde{\mathbb{P}}\{TTT\} = \frac{1}{8}, \\ \widetilde{\mathbb{P}}\{\omega \in \Omega; X(\omega) = 1\} &= \widetilde{\mathbb{P}}\{HTT, THT, TTH\} = \frac{3}{8}, \\ \widetilde{\mathbb{P}}\{\omega \in \Omega; X(\omega) = 2\} &= \widetilde{\mathbb{P}}\{HHT, HTH, THH\} = \frac{3}{8}, \\ \widetilde{\mathbb{P}}\{\omega \in \Omega; X(\omega) = 3\} &= \widetilde{\mathbb{P}}\{HHH\} = \frac{1}{8}. \end{aligned}$$

We shorten the cumbersome notation $\widetilde{\mathbb{P}}\{\omega \in \Omega; X(\omega) = j\}$ to simply $\widetilde{\mathbb{P}}\{X = j\}$. It is helpful to remember, however, that the notation $\widetilde{\mathbb{P}}\{X = j\}$ refers to the probability of a subset of Ω, the set of elements ω for which $X(\omega) = j$. Under $\widetilde{\mathbb{P}}$, the probability that X takes the four values 0, 1, 2, and 3 are

$$\begin{aligned} \widetilde{\mathbb{P}}\{X = 0\} &= \frac{1}{8}, \ \widetilde{\mathbb{P}}\{X = 1\} = \frac{3}{8}, \\ \widetilde{\mathbb{P}}\{X = 2\} &= \frac{3}{8}, \ \widetilde{\mathbb{P}}\{X = 3\} = \frac{1}{8}. \end{aligned}$$

This table of probabilities where X takes its various values records the *distribution* of X under $\widetilde{\mathbb{P}}$.

The random variable Y is different from X because it counts tails rather than heads. However, under $\widetilde{\mathbb{P}}$, the distribution of Y is the same as the distribution of X:

$$\widetilde{\mathbb{P}}\{Y=0\} = \frac{1}{8},\ \widetilde{\mathbb{P}}\{Y=1\} = \frac{3}{8},$$
$$\widetilde{\mathbb{P}}\{Y=2\} = \frac{3}{8},\ \widetilde{\mathbb{P}}\{Y=3\} = \frac{1}{8}.$$

The point here is that the random variable is a function defined on Ω, whereas its distribution is a tabulation of probabilities that the random variable takes various values. A random variable is not a distribution.

Suppose, moreover, that we choose a probability measure $\mathbb{P}$ for Ω that corresponds to a $\frac{2}{3}$ probability of head on each toss and a $\frac{1}{3}$ probability of tail. Then

$$\mathbb{P}\{X=0\} = \frac{1}{27},\ \mathbb{P}\{X=1\} = \frac{6}{27},$$
$$\mathbb{P}\{X=2\} = \frac{12}{27},\ \mathbb{P}\{X=3\} = \frac{8}{27}.$$

The random variable X has a different distribution under $\mathbb{P}$ than under $\widetilde{\mathbb{P}}$. It is the same random variable, counting the total number of heads, regardless of the probability measure used to determine its distribution. This is the situation we encounter later when we consider an asset price under both the actual and the risk-neutral probability measures.

Incidentally, although they have the same distribution under $\widetilde{\mathbb{P}}$, the random variables X and Y have different distributions under $\mathbb{P}$. Indeed,

$$\mathbb{P}\{Y=0\} = \frac{8}{27},\quad \mathbb{P}\{Y=1\} = \frac{12}{27},$$
$$\mathbb{P}\{Y=2\} = \frac{6}{27},\quad \mathbb{P}\{Y=3\} = \frac{1}{27}. \qquad \square$$

Definition 2.2.4. *Let X be a random variable defined on a finite probability space $(\Omega, \mathbb{P})$. The* expectation *(or* expected value*) of X is defined to be*

$$\mathbb{E}X = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega).$$

When we compute the expectation using the risk-neutral probability measure $\widetilde{\mathbb{P}}$, we use the notation

$$\widetilde{\mathbb{E}}X = \sum_{\omega \in \Omega} X(\omega)\widetilde{\mathbb{P}}(\omega).$$

The variance *of X is*

$$\mathrm{Var}(X) = \mathbb{E}\Big[(X - \mathbb{E}X^2\Big].$$

http://www.springer.com/978-0-387-40100-3

Stochastic Calculus for Finance I
The Binomial Asset Pricing Model
Shreve, S.
2004, XV, 187 p., Hardcover
ISBN: 978-0-387-40100-3