1

Probability spaces

In all spheres of life we make decisions based upon incomplete information. Frameworks for predicting the uncertain outcomes of future events have been around for centuries, notably in the age-old pastime of gambling. Much of modern finance draws on this experience. Probabilistic models have become an essential feature of financial market practice.

We begin at the beginning: this chapter is an introduction to basic concepts in probability, motivated by simple models for the evolution of stock prices. Emphasis is placed on the collection of events whose probability we need to study, together with the probability function defined on these events. For this we use the machinery of measure theory, including the construction of Lebesgue measure on $\mathbb{R}$. We introduce and study integration with respect to a measure, with emphasis on powerful limit theorems. In particular, we specialise to the case of Lebesgue integral and compare it with the Riemann integral familiar to students of basic calculus.

1.1 Discrete examples

The crucial feature of financial markets is uncertainty related to the future prices of various quantities, like stock prices, interest rates, foreign exchange rates, market indices, or commodity prices. Our goal is to build a mathematical model capturing this aspect of reality.

Example 1.1
Consider how we could model stock prices. The current stock price (the **spot price**) is usually known, say 10. We may be interested in the price at some fixed future time. This future price involves some uncertainty. Suppose first that in this period of time the stock price jumps a number of times, going either up or down by 0.50 (such a price change is called a **tick**). After two such jumps there will be three possible prices: 9, 10, 11. After 20 jumps there will be a wider range of possible prices: $0, 1, 2, \ldots, 19, 20$.

The set of all possible outcomes will be denoted by Ω and called the **sample space**. The elements of Ω will be denoted by ω. For now we assume that Ω is a finite set.

Example 1.2
If we are interested in the prices after two jumps, we could take $\Omega = \{9, 10, 11\}$. If we want to describe the prices after 20 jumps, we would take $\Omega = \{0, 1, 2, \ldots, 19, 20\}$.

The next step in building a model is to answer to the following question: for a subset $A \subset \Omega$, called an **event**, what is the probability that the outcome lies in A? The number representing the answer will be denoted by $P(A)$, and the convention is to require $P(A) \in [0, 1]$ with $P(\Omega) = 1$ and $P(\varnothing) = 0$. We shall write $p_\omega = P(\{\omega\})$ for any $\omega \in \Omega$. Given p_ω for all $\omega \in \Omega$, the function P is then constructed for any $A \subset \Omega$ by adding the values attached to the elements of A,

$$P(A) = \sum_{\omega \in A} p_\omega.$$

This immediately implies an important property of P, called **additivity**,

$$P(A \cup B) = P(A) + P(B) \quad \text{for any disjoint events } A, B.$$

By induction, it readily extends to

$$P\left(\bigcup_{i=1}^{m} A_i\right) = \sum_{i=1}^{m} P(A_i) \quad \text{for any pairwise disjoint events } A_1, \ldots, A_m.$$

Example 1.3
Consider $\Omega = \{9, 10, 11\}$. The simplest choice is to assign equal probabilities $p_9 = p_{10} = p_{11} = \frac{1}{3}$ to all single-element subsets of Ω.

Example 1.4
In the case of $\Omega = \{0, 1, 2, \dots, 19, 20\}$ we could, once again, try equal probabilities for all single-element subsets of Ω, namely, $p_0 = p_1 = \cdots = p_{20} = \frac{1}{21}$.

The **uniform probability** on a finite Ω assigns equal probabilities $p_\omega = \frac{1}{N}$ for each $\omega \in \Omega$, where N is the number of elements in Ω.

Example 1.5
Uniform probability does not appear to be consistent with the scheme in Example 1.1, where the stock prices result from consecutive jumps by ± 0.50 from an initial price 10. In the case of two consecutive jumps one might argue that the middle price 10 should carry more weight since it can be arrived at in two ways (up–down or down–up), while either of the other two values can occur in just one way (down–down for 9, up–up for 11). Hence, price 10 would be twice as likely as 9 or 11.

To reflect these considerations on $\Omega = \{9, 10, 11\}$ we can take $p_9 = \frac{1}{4}$, $p_{10} = \frac{1}{2}$, $p_{11} = \frac{1}{4}$.

Example 1.6
Similarly, for $\Omega = \{0, 1, 2, \dots, 19, 20\}$ we can take $p_n = \binom{20}{n}\frac{1}{2^{20}}$, where $\binom{20}{n} = \frac{20!}{n!(20-n)!}$ is the number of scenarios consisting of n upwards and $20-n$ downwards price jumps of 0.50 from the initial price 10, with each scenario equally likely. This is illustrated in Figure 1.1.

In general, when for an N-element Ω we have $p_n = \binom{N}{n}\frac{1}{2^N}$, we call this the **symmetric binomial probability**. Clearly, $\sum_{n=0}^{N} p_n = 1$.

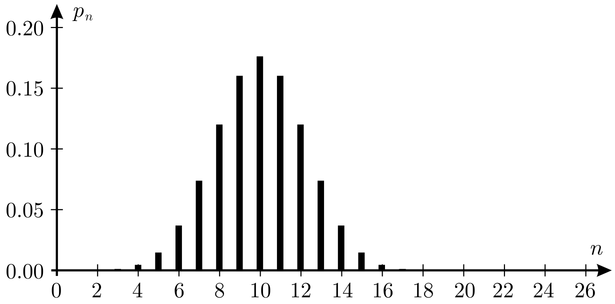

Figure 1.1 Binomial probability and additive jumps.

The mechanism of price jumps by constant additive ticks is not entirely satisfactory as a model for stock prices. After sufficiently many jumps, the range of possible prices will include negative values. To have a more realistic model we need to adjust this mechanism of price jumps.

Example 1.7
The first price jump of ± 0.50 means that the price changes by $\pm 5\%$. In subsequent steps we shall now allow the prices to go up or down by 5% rather than by a constant tick of 0.50. The possible prices will then be $\Omega = \{\omega_n : n = 0, 1, 2, \ldots, 19, 20\}$ after 20 jumps, with $\omega_n = 10 \times 1.05^n \times 0.95^{20-n}$. The prices will remain positive for any number of jumps. We choose the probabilities in a similar manner as before, $p_{\omega_n} = \binom{20}{n}\frac{1}{2^{20}}$. Compare Figure 1.2 with Figure 1.1 to observe a subtle but crucial shift in the distribution of stock prices.

The above examples restrict the possible stock prices to a finite set. In an attempt to extend the model we might want to allow an infinite sequence of possible prices, that is, a countable set Ω.

Example 1.8
Suppose that the number of stock price jumps occurring within a fixed time period is not prescribed, but can be an arbitrary integer N. To be specific, suppose that the probability of N jumps is

$$q_N = \frac{\lambda^N e^{-\lambda}}{N!}$$

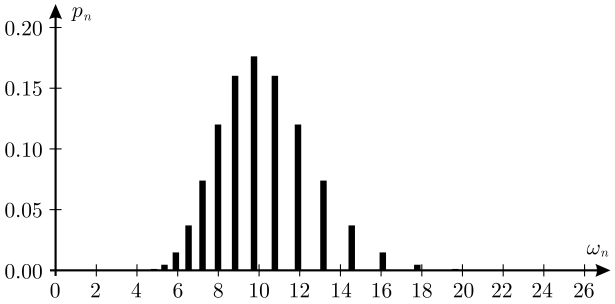

Figure 1.2 Binomial probability and multiplicative jumps.

with $N = 0, 1, 2, \ldots$ for some parameter $\lambda > 0$. The probability of large N is small, but there is no upper bound on N, allowing for some hectic trading. Clearly,

$$\sum_{N=0}^{\infty} q_N = \sum_{N=0}^{\infty} \frac{\lambda^N e^{-\lambda}}{N!} = e^{-\lambda} \sum_{N=0}^{\infty} \frac{\lambda^N}{N!} = e^{-\lambda} e^{\lambda} = 1.$$

This is called the **Poisson probability** with parameter λ.

Furthermore, conditioned on there being N jumps, the possible final stock prices will be described by means of the binomial probability and multiplicative jumps. We assume, like in Example 1.7, that each jump incerases/reduces the stock price by 5% with probability $\frac{1}{2}$. The stock price at time T will become

$$S(T) = 10 \times 1.05^n \times 0.95^{N-n}$$

with probability

$$p_{N,n} = q_N \binom{N}{n} \frac{1}{2^N},$$

that is, the probability q_N of $N = 0, 1, 2, \ldots$ jumps multiplied by the probability $\binom{N}{n}\frac{1}{2^N}$ of n upwards price movements among those N jumps, where $0 \le n \le N$. We take Ω to be the set of such pairs of integers N, n. The formula $P(A) = \sum_{\omega \in A} p_\omega$ defining the probability of an event now includes infinite sets $A \subset \Omega$.

This example shows that it is natural to consider a stronger version of the additivity property:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

for any sequence of pairwise disjoint events $A_1, A_2, \ldots \subset \Omega$. This is known as **countable additivity**.

Example 1.9
Another example where a countable set emerges in a natural way is related to modelling the instant when something unpredictable may happen. Time is measured by the number of discrete steps (of some fixed but unspecified length). At each step there is an upward/downward price jump with probabilities $p, 1-p \in (0,1)$, respectively. The probability that an upward jump occurs for the first time at the nth step can be expressed as $p_n = (1-p)^{n-1}p$. It is easy to check that $\sum_{n=1}^{\infty} p_n = 1$, which gives a probability on $\Omega = \{1, 2, \ldots\}$. This defines the **geometric probability**.

1.2 Probability spaces

Countable additivity turns out to be the perfect condition for probability theory. The actual construction of a probability measure can present difficulties. In particular, it is sometimes impossible to define P for all subsets of Ω. The domain of P has to be specified, and it is natural to impose some restrictions on that domain ensuring that countable additivity can be formulated.

Definition 1.10
A **probability space** is a triple $(\Omega, \mathcal{F}, P)$ as follows.

(i) Ω is a non-empty set (called the **sample space**, or set of scenarios).
(ii) $\mathcal{F}$ is a family of subsets of Ω (called **events**) satisfying the following conditions:
- $\Omega \in \mathcal{F}$;
- if $A_i \in \mathcal{F}$ for $i = 1, 2, \ldots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (we say that $\mathcal{F}$ is **closed under countable unions**);
- if $A \in \mathcal{F}$, then $\Omega \setminus A \in \mathcal{F}$ (we say that $\mathcal{F}$ is **closed under complements**).

Such a family of sets $\mathcal{F}$ is called a **σ-field** on Ω.

(iii) P assigns numbers to events,

$$P : \mathcal{F} \to [0, 1],$$

and we assume that

- $P(\Omega) = 1$;
- for all sequences of events $A_i \in \mathcal{F}$, $i = 1, 2, 3, \ldots$ that are pairwise disjoint ($A_i \cap A_j = \varnothing$ for $i \neq j$) we have

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

This property is called **countable additivity**.

A function P satisfying these conditions is called a **probability measure** (or simply a **probability**).

Exercise 1.1 Let $\mathcal{F}$ be a σ-field and $A_1, A_2, \ldots \in \mathcal{F}$. Show that $\bigcap_{i=1}^{n} A_i \in \mathcal{F}$ for each $n = 1, 2, \ldots$ and that $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

Exercise 1.2 Suppose that $\mathcal{F}$ is a σ-field containing all open intervals in $[0, 1]$ with rational endpoints. Show that $\mathcal{F}$ contains all open intervals in $[0, 1]$.

Before proceeding further we note some basic properties of probability measures.

Theorem 1.11

If P is a probability measure, then:

(i) $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$ *for any pairwise disjoint events* $A_i \in \mathcal{F}$, $i = 1, 2, \ldots, n$ (**finite additivity**)*;*

(ii) $P(\Omega \setminus A) = 1 - P(A)$ *for any* $A \in \mathcal{F}$*; in particular,* $P(\varnothing) = 0$*;*

(iii) $A \subset B$ *implies* $P(A) \leq P(B)$ *for any* $A, B \in \mathcal{F}$ (**monotonicity**)*;*

(iv) $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$ *for any* $A_i \in \mathcal{F}$, $i = 1, 2, \ldots, n$ (**finite subadditivity**)*;*

(v) *if* $A_{n+1} \supset A_n \in \mathcal{F}$ *for all* $n \geq 1$*, then* $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{m\to\infty} P(A_m)$*;*

(vi) $P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ *for any* $A_i \in \mathcal{F}$, $i = 1, 2, \ldots$ (**countable subadditivity**)*;*

(vii) *if* $A_{n+1} \subset A_n \in \mathcal{F}$ *for all* $n \geq 1$*, then* $P(\bigcap_{n=1}^{\infty} A_n) = \lim_{m\to\infty} P(A_m)$.

Proof (i) Let $A_{n+1} = A_{n+2} = \cdots = \varnothing$ and apply countable additivity.

(ii) Use (i) with $n = 2$, $A_1 = A$, $A_2 = \Omega \setminus A$.

(iii) Since $B = A \cup (B \setminus A)$ and we have disjoint components, we can apply (i), so

$$P(B) = P(A) + P(B \setminus A) \geq P(A).$$

(iv) For $n = 2$,

$$P(A_1 \cup A_2) = P(A_1 \cup (A_2 \setminus A_1)) = P(A_1) + P(A_2 \setminus A_1) \leq P(A_1) + P(A_2)$$

and then use induction to complete the proof for arbitrary n, where the induction step will be the same as the above argument.

(v) Using the above properties, we have (with $A_0 = \varnothing$)

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(\bigcup_{n=0}^{\infty} (A_{n+1} \setminus A_n)\right) = \sum_{n=0}^{\infty} P(A_{n+1} \setminus A_n) \\ &= \lim_{m\to\infty} \sum_{n=0}^{m} P(A_{n+1} \setminus A_n) = \lim_{m\to\infty} P\left(\bigcup_{n=0}^{m} (A_{n+1} \setminus A_n)\right) \\ &= \lim_{m\to\infty} P(A_{m+1}). \end{aligned}$$

(vi) We put $B_n = \bigcup_{i=1}^{n} A_i$, so that $B_{n+1} \supset B_n \in \mathcal{F}$ for all $n \geq 1$, and using (v) we pass to the limit in the finite subadditivity relation (iv):

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{m\to\infty} P(B_m) = \lim_{m\to\infty} P\left(\bigcup_{n=1}^{m} A_n\right) \\ &\leq \lim_{m\to\infty} \sum_{n=1}^{m} P(A_n) = \sum_{n=1}^{\infty} P(A_n). \end{aligned}$$

(vii) Take $A = \bigcap_{n=1}^{\infty} A_n$, note that $P(A) = 1 - P(\Omega \setminus A)$ by (ii), and apply (v):

$$\begin{aligned} P(A) &= 1 - P\left(\bigcup_{n=1}^{\infty} (\Omega \setminus A_n)\right) \\ &= 1 - \lim_{m\to\infty} P(\Omega \setminus A_m) = \lim_{m\to\infty} P(A_m). \end{aligned}$$

□

The construction of interesting probability measures requires some labour, as we shall see. However, a few simple examples can be given immediately.

Example 1.12
Take any non-empty set Ω, fix $\omega \in \Omega$, and define $\delta_\omega(A) = 1$ if $\omega \in A$ and $\delta_\omega(A) = 0$ if $\omega \notin A$, for any $A \subset \Omega$. It is a probability measure, called the **unit mass**, also known as the **Dirac measure,** concentrated at ω. If $\mathcal{F}$ is taken to be the family of all subsets of Ω, then $(\Omega, \mathcal{F}, \delta_\omega)$ is a probability space.

Example 1.13
Let N be a positive integer. On $\Omega = \{0, 1, \ldots, N\}$ define

$$P(A) = \sum_{n=0}^{N} \binom{N}{n} \frac{1}{2^N} \delta_n(A)$$

for any $A \subset \Omega$, where δ_n is the unit mass concentrated at n from Example 1.12. We take $\mathcal{F}$ to be the family of all subsets of Ω. Then $(\Omega, \mathcal{F}, P)$ is a probability space. This is clearly the symmetric binomial probability considered earlier.

More generally for the same Ω and any $p \in (0, 1)$, the **binomial probability** with parameters N, p is defined by setting

$$P(A) = \sum_{n=0}^{N} \binom{N}{n} p^n (1-p)^{N-n} \delta_n(A).$$

It is immediate from the binomial theorem that $P(\Omega) = 1$.

This example is often described as providing the probabilities of events relating to the repeated tossing of a coin (where successive tosses are assumed not to affect each other, in a sense that will be made precise later): if for any given toss the probability of 'Heads' is p, the probability of finding exactly k 'Heads' in N tosses is $\binom{N}{k} p^k (1-p)^{N-k}$.

Example 1.14
Fix $\lambda > 0$, let $\Omega = \{0, 1, 2, \ldots\}$ and let $\mathcal{F}$ be the family of all subsets of Ω. For any $A \in \mathcal{F}$ put

$$P(A) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \delta_n(A),$$

where δ_n is the unit mass concentrated at n. Then $(\Omega, \mathcal{F}, P)$ is a probability space. This gives the Poisson probability mentioned in Example 1.8.

In addition to subsets of $\mathbb{R}$, for example the set $[0, \infty)$ of all non-negative real numbers, it often proves convenient to consider sets containing ∞ or $-\infty$ in addition to real numbers. For instance, we write $[-\infty, \infty]$ for the set of all real numbers in $\mathbb{R}$ together with ∞ and $-\infty$, and $[0, \infty]$ to denote the set of all non-negative real numbers together with ∞.

Probability measures belong to a wider class of countably additive set functions taking values in $[0, \infty]$. Let Ω be a non-empty set and let $\mathcal{F}$ be a σ-field of subsets of Ω.

Definition 1.15
We say that $\mu : \mathcal{F} \to [0, \infty]$ is a **measure** and call $(\Omega, \mathcal{F}, \mu)$ a **measure space** if

(i) $\mu(\varnothing) = 0$;
(ii) for any pairwise disjoint sets $A_i \in \mathcal{F}$, $i = 1, 2, \ldots$

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Note that some of the terms $\mu(A_i)$ in the sum may be infinite, and we use the convention $x + \infty = \infty$ for any $x \in [0, \infty]$.

Moreover, we call μ a **finite measure** if, in addition, $\mu(\Omega) < \infty$.

The properties listed in Theorem 1.11 and their proofs can readily be adapted to the case of an arbitrary measure.

Corollary 1.16
Properties (i) *and* (iii)–(vi) *listed in Theorem 1.11 remain true for any measure* μ*. If we assume in addition that* $\mu(\Omega) < \infty$*, then* (ii) *becomes* $\mu(\Omega \setminus A) = \mu(\Omega) - \mu(A)$*. Moreover, if* $\mu(A_1)$ *is finite, then* (vii) *still holds.*

Example 1.17
For any non-empty set Ω and any $A \subset \Omega$ let

$$\mu(A) = \sum_{\omega \in A} \delta_\omega(A),$$

where δ_ω is the unit mass concentrated at ω. The sum is equal to the number