Options and Mathematics: Lecture 19

December 3, 2020

General probability spaces

We consider briefly probability spaces consisting of an infinite number of sample points.

Assume first that Ω is a countable set. This means that

$$\Omega = \{\omega_n\}_{n\in\mathbb{N}}.$$

For countable sample spaces the definitions given for finite sets extend straightforwardly. Precisely, given a sequence

$$p = (p_n)_{n\in\mathbb{N}} \quad \text{such that} \quad 0 < p_n < 1, \quad \sum_{n\in\mathbb{N}} p_n = 1,$$

we define the probability of the atomic events as

$$\mathbb{P}(\{\omega_n\}) = p_n.$$

If $A \in 2^\Omega$, then we define

$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

If $X : \Omega \to \mathbb{R}$ is a random variable, then

$$\mathbb{E}[X] = \sum_{n \in \mathbb{N}} X(\omega_n) p_n = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\{\omega\}).$$

The remaining definitions introduced in the finite case (variance, covariance, independent random variables, etc.) continue to be valid for countable probability spaces.

In the rest of the lecture we assume that Ω is uncountable (e.g., $\Omega = \mathbb{R}$).

In this case there is no general procedure to construct a probability space, but only an abstract definition. In particular a probability measure $\mathbb{P}$ on events $A \subseteq \Omega$ is defined only axiomatically by requiring that

$$0 \leq \mathbb{P}(A) \leq 1, \quad \mathbb{P}(\emptyset) = 0$$

and that, for any sequence of disjoint events $A_1, A_2, \dots$, it should hold

$$\mathbb{P}(A_1 \cup A_2 \cup \dots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \dots$$

Moreover it is not necessary—and almost never convenient—to assume that $\mathbb{P}$ is defined for all events $A \subset \Omega$.

We denote by $\mathcal{F}$ the set of events (i.e., subsets of Ω) which have a well defined probability satisfying the properties above.

Example.

Let $\Omega = \mathbb{R}$. We say that $A \subseteq \mathbb{R}$ is a **Borel set** if it can be written as the union (or intersection) of countably many open (or closed) intervals.

Let $\mathcal{F}$ be the collection of all Borel sets. Given a continuous, non-negative function $p : \mathbb{R} \to [0, \infty)$ such that

$$\int_{\mathbb{R}} p(x)\, dx = 1,$$

we define $\mathbb{P} : \mathcal{F} \to [0, 1]$ as

$$\mathbb{P}(A) = \int_A p(\omega)\, d\omega.$$

If $X : \mathbb{R} \to \mathbb{R}$ is a random variable, the expectation of X in this probability measure is given by

$$\mathbb{E}[X] = \int_{\mathbb{R}} X(\omega) p(\omega)\, d\omega,$$

provided the integral is finite.

For most applications the knowledge of the full probability space is not necessary.

We are only interested in assigning a probability to events of the form

$$\{X \in I\}$$

where X is a random variable on the sample space Ω and $I \subset \mathbb{R}$ is a Borel set (e.g., an interval), that is to say, events which can be resolved by one (or more) random variables.

The probability $\mathbb{P}(X \in I)$ can be computed explicitly when X has a density.

Definition 6.1

A random variable $X : \Omega \to \mathbb{R}$ is said to have **probability density** f_X if $f_X : \mathbb{R} \to [0, \infty)$ is continuous and

$$\mathbb{P}(X \in A) = \int_A f_X(x)\, dx,$$

for all Borel sets $A \subseteq \mathbb{R}$.

Note that the density f_X satisfies

$$\int_{\mathbb{R}} f_X(x)\, dx = 1$$

and the cumulative distribution $F_X(x) = \mathbb{P}(X \le x)$ satisfies

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy, \quad \text{for all } x \in \mathbb{R}, \text{ hence } f_X = \frac{dF_X}{dx}.$$

Example:

A random variable $X : \Omega \to \mathbb{R}$ is said to be a **normal random variable** with mean $m \in \mathbb{R}$ and variance $\sigma^2 > 0$ if it admits the density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|x - m|^2}{2\sigma^2}\right).$$

We denote by $\mathcal{N}(m, \sigma^2)$ the set of all such random variables. A variable $X \in \mathcal{N}(0, 1)$ is called a **standard normal random variable**. The cumulative distribution of standard normal random variables is denoted by $\Phi(x)$ and is called the **standard normal distribution**, i.e.,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2}\, dy.$$

Moreover we denote by $\phi(x)$ the **standard normal (probability) density**, that is

$$\phi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Theorem 6.1

The following holds for all sufficiently regular functions $g : \mathbb{R} \to \mathbb{R}$:

(i) Let $X : \Omega \to \mathbb{R}$ be a random variable with density f_X. Then for all Borel sets $A \subseteq \mathbb{R}$,

$$\mathbb{P}(g(X) \in A) = \int_{x:g(x)\in A} f_X(x)\, dx,$$

(ii) Let $X : \Omega \to \mathbb{R}$ be a random variable with density f_X. Then

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(y) f_X(y)\, dy,$$

Moreover the properties 1,2,3 in Theorem 5.2 still hold for continuum random variables.

Example

As an example of application of (i), suppose $X \in \mathcal{N}(0,1)$. Then

$$\mathbb{P}(X^2 \le 1) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{x^2}{2}}\, dx \approx 0.683,$$

that is to say, a standard normal random variable has about 68,3% probability to take value on the interval $(-1, 1)$.

By (ii) in Theorem 6.1, the expectation and the variance of a continuum random variable X with density f_X are given by

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)\, dx, \quad \mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_{\mathbb{R}} x^2 f_X(x)\, dx - \left(\int_{\mathbb{R}} x f_X(x)\, dx \right)^2.$$

In particular, for normal variables we obtain

$$X \in \mathcal{N}(m, \sigma^2) \implies \mathbb{E}[X] = m, \quad \mathrm{Var}[X] = \sigma^2.$$

Stochastic processes. Martingales

Let $\{X(t)\}_{t\geq 0}$ be a stochastic process, i.e., a one parameter family of (continuum) random variables and denote by $\mathcal{F}_X(t)$ the information accumulated by "looking" at the stochastic process up to time t, i.e., the collection of events resolved by $X(s)$ for $0 \leq s \leq t$.

Intuitively, the stochastic process $\{X(t)\}_{t\geq 0}$ is a martingale if, based on the information contained in $\mathcal{F}_X(s)$, our "best estimate" on $X(t)$ for $t > s$ is $X(s)$, i.e., we are not able to estimate whether the process will raise or fall in the interval $[s, t]$ with the information available at time s.

This intuitive definition is encoded in the formula

$$\mathbb{E}[X(t)|\mathcal{F}_X(s)] = X(s), \quad 0 \leq s \leq t,$$

which generalizes the definition of martingales in finite probability theory.

The left hand side of the previous formula is the conditional expectation of $X(t)$ with respect to the information $\mathcal{F}_X(s)$, whose precise definition is not needed here.

It can be shown that martingales have constant expectation.

Brownian motion

Next we define the most important of all stochastic processes.

Definition 6.3

A **Brownian motion**, or **Wiener process**, is a stochastic process $\{W(t)\}_{t\geq 0}$ with the following properties:

1. For all $\omega \in \Omega$, the paths are continuous (i.e., $t \to W(t,\omega)$ is a continuous function) and $W(0,\omega) = 0$;
2. For all $0 = t_0 < t_1 < t_2 < \dots$, the increments
$$W(t_1) = W(t_1) - W(t_0),\ W(t_2) - W(t_1), \dots,$$
are independent random variables and
$$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0, \quad \mathrm{Var}[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i, \quad \text{for all } i = 0, 1, \dots;$$
3. The increments are normally distributed, that is to say, for all $0 \leq s < t$,
$$\mathbb{P}(W(t) - W(s) \in A) = \frac{1}{\sqrt{2\pi(t-s)}} \int_A e^{-\frac{y^2}{2(t-s)}}\, dy,$$
for all Borel sets $A \subseteq \mathbb{R}$.

It can be shown that Brownian motions exist, yet a formal construction is technically quite difficult and beyond the purpose of this course

It is useful to think of Brownian motions as time-continuum generalizations of the symmetric random walk.

Remarks.

- Since the definition of Brownian motion depends on the probability measure $\mathbb{P}$, then a stochastic process $\{W(t)\}_{t\geq 0}$ which is a Brownian motion in the probability measure $\mathbb{P}$ will in general *not* be a Brownian motion in another probability measure $\widetilde{\mathbb{P}}$. When we want to emphasize that $\{W(t)\}_{t\geq 0}$ is a Brownian motion in the probability measure $\mathbb{P}$, we shall say that $\{W(t)\}_{t\geq 0}$ is a **$\mathbb{P}$-Brownian motion.**

- Letting $s = 0$ in property 3 in Definition 6.3 we obtain that $W(t) \in \mathcal{N}(0,t)$, for all $t > 0$. In particular, $W(t)$ has zero expectation for all times. It can also be shown that Brownian motions are martingales.

Equivalent probability measures. Girsanov theorem

One further technical complication arising for uncountable sample spaces is the existence of non-trivial events with zero probability, e.g., the event $\{W(1) = 0\}$.

We shall need to consider the concept of equivalent probability measures.

Definition 6.4

Two probability measure $\mathbb{P}$, $\widetilde{\mathbb{P}}$ are said to be equivalent if $\mathbb{P}(A) = 0 \Leftrightarrow \widetilde{\mathbb{P}}(A) = 0$ for all events $A \in \mathcal{F}$.

Hence equivalent probability measures agree on which events are impossible.

Note that in a finite probability space all probability measures are equivalent, as in the finite case the empty set is the only event with zero probability.

The main question we now want to answer is the following: Given a probability measure $\mathbb{P}$, how can we find all probability measures $\widetilde{\mathbb{P}}$ equivalent to $\widetilde{\mathbb{P}}$?

The answer to this question is given in the Radon-Nikodým theorem.

Theorem 6.7 (Radon-Nikodým theorem)

Let $\mathbb{P} : \mathcal{F} \to [0,1]$ be a probability measure. Then $\widetilde{\mathbb{P}} : \mathcal{F} \to [0,1]$ is a probability measure equivalent to $\mathbb{P}$ if and only if there exists a random variable $Z : \Omega \to \mathbb{R}$ such that $Z > 0$ (with probability 1), $\mathbb{E}[Z] = 1$ and

$$\widetilde{\mathbb{P}}(A) = \mathbb{E}[Z\mathbb{I}_A], \quad \text{for all} \quad A \in \mathcal{F}, \quad \text{where} \quad \mathbb{I}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}.$$

Examples

- Assume $\Omega = \mathbb{R}$ and that $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are defined as in the example on page 3, namely
$$\mathbb{P}(A) = \int_A p(\omega)\, d\omega, \quad \widetilde{\mathbb{P}}(A) = \int_A \widetilde{p}(\omega)\, d\omega,$$
for all Borel sets A and for some given continuous non-negative functions p, $\widetilde{p}$ such that
$$\int_{\mathbb{R}} p(\omega)\, d\omega = \int_{\mathbb{R}} \widetilde{p}(\omega)\, d\omega = 1.$$
Then, according to Theorem 6.7, $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are equivalent if and only if there exists a function $Z : \mathbb{R} \to \mathbb{R}$ such that $Z > 0$ and
$$\widetilde{\mathbb{P}}(A) = \int_A \widetilde{p}(\omega)\, d\omega = \int_{\mathbb{R}} Z(\omega)\mathbb{I}_A(\omega)p(\omega)\, d\omega$$
that is
$$\int_A \widetilde{p}(\omega)\, d\omega = \int_A Z(\omega)p(\omega)\, d\omega$$
It can be shown that the latter equality is satisfied for all Borel sets $A \subset \mathbb{R}$ if and only if $\widetilde{p}(\omega) = Z(\omega)p(\omega)$ (with probability 1).

- Let $\{W(t)\}_{t\geq 0}$ be a $\mathbb{P}$-Brownian motion, $\theta \in \mathbb{R}$ and $T > 0$. Define

  $$Z_\theta = e^{-\theta W(T) - \frac{1}{2}\theta^2 T}.$$

  Clearly $Z_\theta > 0$ and moreover

  $$\mathbb{E}[Z_\theta] = \mathbb{E}[e^{-\theta W(T) - \frac{1}{2}\theta^2 T}] = \int_{\mathbb{R}} e^{-\theta x - \frac{1}{2}\theta^2 T} \frac{e^{-\frac{x^2}{2T}}}{\sqrt{2\pi T}}\, dx = 1,$$

  where we used the density of the normal random variable $W(T) \in \mathcal{N}(0, T)$ to compute the expectation of Z_θ in the probability measure $\mathbb{P}$. Thus, according to the Radon-Nikodým theorem, $\mathbb{P}_\theta(A) = \mathbb{E}[Z_\theta \mathbb{I}_A]$ is a probability measure equivalent to $\mathbb{P}$, for all $\theta \in \mathbb{R}$.

As the last example will play a crucial role in the discussion on Black-Scholes theory, it deserves a proper definition.

Definition 6.5

Let $\{W(t)\}_{t\geq 0}$ be a $\mathbb{P}$-Brownian motion, $\theta \in \mathbb{R}$, $T > 0$ and Z_θ be the random variable

$$Z_\theta = e^{-\theta W(T) - \frac{1}{2}\theta^2 T}.$$

The probability measure $\mathbb{P}_\theta(A) = \mathbb{E}[Z_\theta \mathbb{I}_A]$ is called **Girsanov probability** with parameter θ.

Note that the Girsanov probability measure $\mathbb{P}_\theta$ depends also on T, but this is not reflected in our notation.

In the following we denote by $\mathbb{E}_\theta[\cdot]$ the expectation computed in the probability measure $\mathbb{P}_\theta$. When $\theta = 0$ we have $\mathbb{P}_0 = \mathbb{P}$, in which case we denote the expectation as usual by $\mathbb{E}[\cdot]$.

Now we can state a fundamental theorem in probability theory with deep applications in financial mathematics, namely Girsanov's theorem.

Theorem 6.7 (Girsanov theorem)

Let $\{W(t)\}_{t\geq 0}$ be a $\mathbb{P}$-Brownian motion. Given $\theta \in \mathbb{R}$, define the stochastic process $\{W^{(\theta)}(t)\}_{t\geq 0}$ by

$$W^{(\theta)}(t) = W(t) + \theta t.$$

Then $\{W^{(\theta)}(t)\}_{t\geq 0}$ is a $\mathbb{P}_\theta$-Brownian motion.

Note carefully that $\{W^{(\theta)}(t)\}_{t\geq 0}$ is *not* a $\mathbb{P}$-Brownian motion for $\theta \neq 0$, because $\mathbb{E}[W^{(\theta)}(t)] = \theta t \neq 0$.

In particular, according to the probability measure $\mathbb{P}$, the stochastic process $\{W^{(\theta)}(t)\}_{t\geq 0}$ has a *drift*, i.e., a tendency to move up (if $\theta > 0$) or down (if $\theta < 0$). However in the Girsanov probability $\mathbb{P}_\theta$ this drift is removed, because $\mathbb{E}_\theta[W^{(\theta)}(t)] = 0$.