622: Lecture 3
Probability Spaces, Random Variables, and Derivative Pricing

Daniel Ocone

September 13, 2011

Probability Spaces

Elements of a probability space:

- An outcome space Ω;
- A collection $\mathcal{F}$ of subsets of Ω: members of $\mathcal{F}$ are called events; $A \in \mathcal{F}$ is the event that the outcome is in A;
- A function $\mathbf{P}$ which assigns to each event $A \in \mathcal{F}$ a probability $\mathbf{P}(A)$.

Example: (Coin toss, one-period binomial model)

$$\Omega = \{-1, 1\}, \quad \mathcal{F} = \Big\{ \{-1\}, \{1\}, \emptyset, \{-1, 1\} \Big\} \text{ (all subsets of } \Omega\text{).}$$

$$\mathbf{P}(\{1\}) = p, \quad \mathbf{P}(\{-1\}) = q = 1 - p, \quad \mathbf{P}(\Omega) = 1, \quad \mathbf{P}(\emptyset) = 0.$$

$$\text{Note:} \qquad \mathbf{P}(\{\omega\}) = q^{(1-\omega)/2} p^{(1+\omega)/2}$$

Probability Spaces: Axioms

Axiom on $\mathcal{F}$:

$$\mathcal{F} \text{ is a } \sigma\text{-algebra}$$

This means:

(i) $\Omega \in \mathcal{F}$

(ii) If $A \in \mathcal{F}$, so is $A^c = \Omega - A$.

(iii) If $A_1, A_2, \cdots$ belong to $\mathcal{F}$, so does $\bigcup_1^\infty A_n$ and also $\bigcap_1^\infty A_n$.

If $\mathcal{F}$ is a σ-algebra, if $A_1, \ldots, A_n$ belong to $\mathcal{F}$, so do $\bigcup_1^n A_i$ and $\bigcap_1^n A_i$.

Probability Spaces: examples of σ-algebras.

Example 1: The collection of all subsets of Ω is a σ algebra, no matter what Ω is.

Example 2: Let $\Omega = \{1,2,3,4,5,6\}$. Let

- $\mathcal{C} = \Big\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\},\Omega,\emptyset\Big\}$
- $\mathcal{G} = \Big\{\{1,2,3\},\{4,5,6\},\Omega,\emptyset\Big\}$

• $\mathcal{C}$ is **not** a σ-algebra; it is not closed under finite unions (eg. $\{1\} \cup \{2\} = \{1,2\} \notin \mathcal{C}$); it is not closed under complementation (eg. $\{1\}^c = \{1\}^c = \{2,3,4,5,6\} \notin \mathcal{C}$.
• $\mathcal{G}$ is a σ-algebra.

Examples of σ-algebras, continued

Example 3: Let $A_1, A_2, \ldots, A_n$ be disjoint and $\bigcup_1^n A_i = \Omega$. Let $\mathcal{F}$ be the collection which contains the empty set and every set which is a union of a sub-collection of $\{A_1, \ldots, A_n\}$.

Thus, $\mathcal{G} = \Big\{\emptyset, A_1, \ldots, A_n, A_1 \cup A_2, \ldots, A_i \cup A_j, \ldots, A_j \cup A_j \cup A_k, \cdots, \Omega\}$.

Example 4: Let $\Omega = [0,1]$. Let $\mathcal{G}$ be the collection of all finite or countable unions of subintervals of $[0,1]$. This is **not** a σ-algebra. This set is not closed under taking complements.

Example 5: Given any collection $\mathcal{G}$ of subsets of Ω, there is always a smallest σ-algebra of subset containing $\mathcal{G}$. It is denoted $\sigma(\mathcal{G})$ and called the σ-algebra generated by $\mathcal{G}$.

Axioms on $\mathbf{P}$

$\mathbf{P}$ must satisfy

(i) For every $A \in \mathcal{F}$, $0 \leq \mathbf{P}(A) \leq 1$.

(ii) $\mathbf{P}(\Omega) = 1$.

(iii) (Countable additivity) If $A_1, A_2, \ldots, A_n, \ldots$ belong to $\mathcal{F}$ and are disjoint then

$$\mathbf{P}\Big(\bigcup_{1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbf{P}(A_i)$$

Note: (iii) implies that if $A_1, \ldots, A_n$ are disjoint events in $\mathcal{F}$,

$$\mathbf{P}\Big(\bigcup_{1}^{n} A_i\Big) = \sum_{i=1}^{n} \mathbf{P}(A_i)$$

Probability Space: Examples

Definition

$(\Omega, \mathcal{F}, \mathbf{P})$ is called a *probability space* if $\mathcal{F}$ is a σ-algebra and $\mathbf{P}$ satisfies axioms (i),(ii),(iii).

Example 1: Roll of a die; $\Omega = \{1, 2, 3, 4, 5, 6\}$;
$\mathcal{F}$ is all subsets of Ω;
$p_1, \ldots, p_6$ are non-negative numbers such that $p_1 + \cdots + p_6 = 1$, and

$$\mathbf{P}(\{k\}) = p_k, \quad 1 \leq k \leq 6,$$

$$\text{for any } A \subset \Omega, \quad \mathbf{P}(A) = \sum_{k \in A} p_k$$

Example 2: (General discrete space)

- $\Omega = \{\omega_1, \omega_2, \dots\}$
- $\mathcal{F}$ is all subsets of Ω
- $p_1, p_2, \dots$ are non-negative and $\sum_1^\infty p_i = 1$.

$$\mathbf{P}(\{\omega_i\}) = p_i \quad \text{for each } i,$$

$$\text{for any } A \subset \Omega, \ \ \mathbf{P}(A) = \sum_{\omega_k \in A} p_k$$

Why not always take $\mathcal{F}$ to be all subsets of Ω?
Isn't it enough to assign probabilities to all individual outcomes, as in the previous examples?
Example 3. (Uniform probability measure on $[0, 1]$.)

- $\Omega = [0, 1]$
- $\mathcal{F}$ is the smallest σ-algebra containing all subintervals of $[0, 1]$. (the Borel σ-algebra)
- Theorem: There is a unique probability measure **P** on $\mathcal{F}$ such that

$$\mathbf{P}((a, b)) = b - a, \quad \text{for any } 0 \leq a < b \leq 1.$$

Notes: 1. $\mathbf{P}(\{x\}) = 0$ for every $x \in [0, 1]$. $\mathbf{P}(A) = 0$ for any countable set $(A = \{x_1, x_2, \ldots\})$. It is only meaningful to assign positive probabilities to events with uncountably many elements.

2. $\mathcal{F}$ is not the same as the collection of all subsets of $[0, 1]$ and it is not possible to extend **P** to this collection so that it satisfies the additivity axiom.

Independence

Definition

Events $A_1, \ldots, A_n$ are independent if for any $1 \leq i_1 < i_2 < \cdots < i_r \leq n$,

$$\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_r}) = \mathbf{P}(A_{i_1}) \cdots \mathbf{P}(A_{i_r})$$

Multi-period market example.

- $\Omega = \Big\{\omega = (\omega_1, \ldots, \omega_n); , \omega_i$ is -1 or 1 for each $i\Big\}$;
- $\mathcal{F}$ is the collection of all subsets of Ω;
- Want: (i) market movement in different periods to be independent.
  (ii) Fix $0 < p < 1$. Probability of $\omega_i = 1$ ("up") for each i equals p.

$q = 1 - p$.

Thus $\mathbf{P}\Big(\{\omega; \omega_i = x_i\}\Big) = q^{(1-x_i)/2} p^{(1+x_i)/2}$, and

$$\begin{aligned}
\mathbf{P}\Big(\{\omega = (x_1, \ldots, x_n)\}\Big) &= \prod_1^n \mathbf{P}\Big(\{\omega_i = x_i\}\Big) = \prod_1^n q^{(1-x_i)/2} p^{(1+x_i)/2} \\
&= q^{(1/2)\sum_1^n (1-x_i)} p^{(1/2)\sum_1^n (1+x_i)}
\end{aligned}$$

For any $A \subset \Omega$,

$$\mathbf{P}(A) = \sum_{(\omega_1, \ldots, \omega_n) \in A} q^{(1/2)\sum_1^n (1-\omega_i)} p^{(1/2)\sum_1^n (1+\omega_i)}$$

Multi-period model

This is a probability model for independent movements in each period of a multi-period, binomial model.

Note: In the formula

$$\mathbf{P}\Big(\{(\omega_1,\ldots,\omega_n)\}\Big)=q^{(1/2)\sum_1^n(1-\omega_i)}p^{(1/2)\sum_1^n(1+\omega_i)}$$

$$N(\omega)=\frac{1}{2}\sum_1^n(1+\omega_i)=\text{ number of 1's in }(\omega_1,\ldots,\omega_n)$$

$$n-N(\omega)=\frac{1}{2}\sum_1^n(1-\omega_i)=\text{ number of }-1\text{'s in }(\omega_1,\ldots,\omega_n)$$

Can write instead $\quad \mathbf{P}\Big(\{(\omega_1,\ldots,\omega_n)\}\Big)=q^{n-N(\omega)}p^{N(\omega)}$

Random variables: Orientation

In elementary probability, a random variable X is treated as a symbol for the numerical outcome of a random experiment or phenomenon, and specified by it probability mass function (when X is discrete) or its cumulative distribution function or density function (when X is continuous.)

In the measure-theoretic approach to probability adhered to in this course, a random variable is *always a function on a probability space*. A probability space is always used as the underlying model. A random variable X assigns to each possible outcome ω a value $X(\omega)$. Since the ω that actually occurs in a given trial is random, the value of X is random. Not any function qualifies as a random variable, as we shall see....

Multi-period, binomial model; examples.

Even though we have not yet defined what qualifies a function as a random variable, we give an important example:

Multi-period binomial model: $\Omega = \Big\{(\omega_1, \ldots, \omega_n);\, \omega_i \in \{-1, 1\}\Big\}$,
$\mathcal{F} =$ all subsets.

Random variables of interest:

- Number of 1's in first k periods: $N_k(\omega_1, \ldots, \omega_n) := \dfrac{1}{2}\sum_1^k (1 + \omega_i)$.
- Number of -1's in first k periods: $k - N_k(\omega) = \dfrac{1}{2}\sum_1^k (1 - \omega_i)$
- Prices (we derived this formula in a previous lecture):

$$S(t_k)(\omega) = u^{N_k(\omega)} d^{k - N_k(\omega)} S(0)$$

(in this formula we abbreviate $(\omega_1, \ldots, \omega_n)$ by ω)

- The payoff of a contingent claim at T: $V(T)(\omega)$, $\omega \in \Omega$.

Random variables; further orientation.

What restrictions should qualify X as a random variable? Recall X is a function on a set Ω which comes from a probability $(\Omega, \mathcal{F}, \mathbf{P})$.

Suppose U is a subset of the real line; for example an interval $[a, b]$, or $(-\infty, b]$. We want to discuss the probability that the value of X follows in an interval. This is the probability of the event

$$\left\{\omega;\ X(\omega) \in U\right\}$$

But to discuss the probability of this set, it is necessary that

$$\left\{\omega;\ X(\omega) \in U\right\} \in \mathcal{F}.$$

Random variables; measurable functions

Let $\mathcal{F}$ be a σ-algebra of subsets of Ω.

Definition

A real-valued function X on Ω is $\mathcal{F}$-measurable (or measurable with respect to $\mathcal{F}$) if

$$\Big\{\omega;\ X(\omega) \leq a\Big\} \in \mathcal{F} \ \text{ for all real numbers } a.$$

Examples:

1. If $\mathcal{F}$ is the collection of all subsets of Ω, all real-valued functions X on Ω are $\mathcal{F}$-measurable.

2. Let $\Omega = \{1,2,3,4,5,6\}$, $\mathcal{F} = \Big\{\{1,3,5\}, \{2,4,6\}, \emptyset, \Omega\Big\}$.
Let $X(j) = j$. $\Big\{\omega; X(\omega) \leq 1\Big\} = \{1\} \notin \mathcal{F}$; thus X is not $\mathcal{F}$-measurable.

More discussion of measurability

Definition

Let $\mathcal{B}(\mathbb{R})$ denote the smallest σ-algebra of subsets of the real line that contains all open intervals. This is called the *Borel σ-algebra* of $\mathbb{R}$ and it elements are called *Borel sets*.

The following are Borel sets: any open set of $\mathbb{R}$, any closed set, any union or intersection of a countable number of open or closed sets, any complement of these, etc. The Borel sets will contain almost any subset that will arise in practice.

Lemma

If X is $\mathcal{F}$-measurable, then

$$\Big\{\omega;\ X(\omega) \in U\Big\} \in \mathcal{F} \quad \textit{for any Borel set } U.$$

Random variables; Definition

Definition

If $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and if X is an $\mathcal{F}$-measurable function defined on Ω, we call X a random variable.

Notes: 1. When Ω is discrete and $\mathcal{F}$ is all subsets of Ω, any function X on Ω qualifies as a random variable.

2. In more complicated examples, measurability can be tricky to check and we will normally assume it is the case.

3. The concept of measurability will be an important component of the theory of conditional probability and expectation.

4. Linear combinations of random variable, limits of sequences of random variables, etc. transformations $y = \psi(X)$ by continuous or piecewise-continuous ψ are all again random variables.

Random variables; simple random variables

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space.

- If $A \subset \Omega$, the indicator function of A is

$$\mathbf{1}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A; \\ 0, & \text{if } \omega \notin A. \end{cases}$$

If $A \in \mathcal{F}$, $\mathbf{1}_A$ is a random variable.

- If $c_1, c_2, \ldots$ are real numbers and if $A_1, A_2, \ldots$ are disjoint events in $\mathcal{F}$, then

$$X = \sum_{i=1}^{\infty} c_i \mathbf{1}_{A_i} \quad \text{is a random variable.}$$

Random variable of this form are called *discrete* random variables. The sum may be finite:

$$X = \sum_{i=1}^{K} c_i \mathbf{1}_{A_i}$$

X takes values only in the set $\{c_1, c_2, \ldots\}$ and $\mathbf{P}(X = c_i) = \mathbf{P}(A_i)$.

Conversely, assume X is a random variable whose possible values lie in the set of distinct numbers, $\{c_1, c_2, \ldots\}$. Let $A_i = \big\{\omega;\ X(\omega) = c_i\big\}$, for each i. Then

$$X = \sum_{i=1}^{\infty} c_i \mathbf{1}_{A_i}.$$

All random variables on a discrete Ω are discrete.

One period model with probabilities.

One period model: $\Omega = \{-1, 1\}$. Assume $\mathbf{P}(\{1\}) = p$, $\mathbf{P}(\{-1\}) = q = 1 - p$.

Let $S(T)(1) = uS(0)$, $S(T)(-1) = dS(0)$, or

$$S(T)(\omega) = uS(0)\mathbf{1}_{\{1\}}(\omega) + dS(0)\mathbf{1}_{\{-1\}}(\omega)$$

A contingent claim: $\quad V(T)(\omega) = V(T)(1)\mathbf{1}_{\{1\}}(\omega) + V(T)(-1)\mathbf{1}_{\{-1\}}(\omega)$

Example for the multi-period model with independent market movements

Back to the multi-period model: $\Omega = \Big\{(\omega_1, \ldots, \omega_n);\ \omega_i \in \{-1, 1\}\Big\}$,

$$\mathcal{F} = \text{all subsets.}$$

$$\mathbf{P}\Big(\{(\omega_1, \ldots, \omega_n)\}\Big) = q^{n-N(\omega)} p^{N(\omega)}$$

N is the total number of 1's in n independent trials with p the probability of a 1 in each trial. Thus

$$N \text{ is binomial random variable with parameters } p \text{ and } n:$$

$$\mathbf{P}(N = j) = \binom{n}{j} p^j q^{n-j} \qquad 0 \leq j \leq n.$$

$\{N = j\}$ consists of the $\binom{n}{j}$ sequences $(\omega_1, \ldots, \omega_n)$ with exactly j 1's. Each sequence in the set has probability $p^j q^{n-j}$.

Expectation of simple random variables

Definition

Let $X = \sum_{i=1}^{\infty} c_i \mathbf{1}_{A_i}$, The integral of X with respect to $\mathbf{P}$ is defined to be

$$\int_\Omega X(\omega) d\mathbf{P}(\omega) = \int_\Omega \sum_{i=1}^{\infty} c_i \mathbf{1}_{A_i} d\mathbf{P}(\omega) = \sum_{i=1}^{\infty} c_i \mathbf{P}(A_i)$$

if the sum makes sense. This integral is called the expected value of X (w.r.t. $\mathbf{P}$) and written $E[X]$ or, $E_{\mathbf{P}}$ when it is important to emphasize the probability measure.

This definition coincides with the definition from elementary probability; if c_i are distinct, $E[X] = \sum_1^{\infty} c_i \mathbf{P}(X_i = c_i)$.

Example; One-period model

One period model: $\Omega = \{-1, 1\}$, with
$\mathbf{P}(\{1\}) = p, \quad \mathbf{P}(\{-1\}) = q = 1 - p,$
$S(T)(\omega) = uS(0)\mathbf{1}_{\{1\}}(\omega) + dS(0)\mathbf{1}_{\{-1\}}(\omega).$
Then

$$E[S(T)] = E[S(T)(-1)\mathbf{1}_{\{-1\}} + S(T)(1)\mathbf{1}_{\{1\}}] = qdS(0) + puS(0).$$

For a contingent claim,

$$E[V(T)] = qV(T)(-1) + pV(T)(1).$$

One period model; Continued

Let $d < 1 + rT < u$. Then there is a price vector

$$\frac{1}{1+rT}(\tilde{q}, \tilde{p}) = \frac{1}{1+rT}\Big(\frac{u-(1+rT)}{u-d}, \frac{1+rT-d}{u-d}\Big)$$

We know $\tilde{q} + \tilde{p} = 1$. Let $\tilde{\mathbf{P}}$ be the measure defined by

$$\tilde{\mathbf{P}}(\{1\}) = \tilde{p}, \quad \tilde{\mathbf{P}}(\{-1\}) = \tilde{q}.$$

Let $\tilde{E}$ denote $E_{\tilde{\mathbf{P}}}$. Then

$$S(0) = \frac{1}{1+rT}\big[\tilde{q}S(T)(-1) + \tilde{p}S(T)(1)\big] = \frac{1}{1+rT}\tilde{E}[S(T)]$$

and, for any contingent claim, the no-arbitrage price is

$$V(0) = \frac{1}{1+rT}\big[\tilde{q}V(T)(-1) + \tilde{p}V(T)(1)\big] = \frac{1}{1+rT}\tilde{E}[V(T)].$$

Probabilistic definition of the state-price vector.

Definition

Consider a one-period model with $\Omega = \{\omega_1, \ldots, \omega_m\}$, risky assets $S_1(t), \ldots, S_p(t)$, and a money market at risk-free rate r. A measure $\mathbf{P}$ on Ω is said to be risk-neutral if $\mathbf{P}(\{\omega_j\}) > 0$ for each j, and if for *each* risky asset i, $1 \leq i \leq p$,

$$S_i(0) = \frac{1}{1+rT} E_{\mathbf{P}}[S_i(T)] = \frac{1}{1+rT}\Big[\sum_{i=1}^{m} S_i(T)(\omega_i)\mathbf{P}\big(\{\omega_i\}\big)\Big].$$

Usually we use $\tilde{\mathbf{P}}$ to denote a risk-neutral measure.

Risk-neutral measures and state-price vectors

Theorem

If $\tilde{\mathbf{P}}$ is a risk-neutral measure,

$$\frac{1}{1+rT}\Big(\tilde{\mathbf{P}}(\{\omega_1\}),\ldots,\tilde{\mathbf{P}}(\{\omega_m\})\Big).$$

is a state-price vector. Conversely, a state-price vector defines a risk-neutral measure.

Fundamental Theorem of Asset Pricing, restated

Theorem

The one-period model is arbitrage-free if and only if there exists a risk-neutral measure $\tilde{\mathbf{P}}$. In this case, if $V(T)$ is an attainable contingent claim, its unique, no-arbitrage price is

$$V(0) = \frac{1}{1+rT}\tilde{E}[V(T)].$$

Assume $d < 1 + rh < u$, where h is the length of each period. Let $\tilde{p}$ and $\tilde{q}$ be defined as in the one-period model. We showed that the value of a contingent claim at time 0 is

$$V(0) = \frac{1}{(1+rh)^n} \sum_{\omega \in \Omega} \tilde{p}^N(\omega) \tilde{q}^{n-N(\omega)} V(T)(\omega).$$

If we define $\tilde{\mathbf{P}}$ on Ω by

$$\tilde{\mathbf{P}}\Big(\{(\omega_1, \ldots, \omega_n)\}\Big) = \tilde{q}^{n-N(\omega)} \tilde{p}^{N(\omega)},$$

it follows

$$V(0) = \frac{1}{(1+rh)^n} \tilde{E}[V(T)].$$

$\tilde{\mathbf{P}}$ will again turn out to be a risk-neutral measure, but we have yet to define this for the multi-period case.