# Introduction

Olivier Scaillet

University of Geneva and Swiss Finance Institute

# Outline

# Course objective

The objective of the course is to introduce stochastic processes and their applications in finance and insurances modelling.

The course starts with a review of probability theory.

# Syllabus

The first class of processes we will study is the class of Markov chains.

We will then consider Poisson and Point processes before stuyding the martingale theory.

The last part of the course will cover diffusion processes.

# Applications

Concrete applications:

- modeling asset prices
- valuation of financial assets
- risk management in finance and insurance

# Justification

« Future is mostly unpredictable »

To express this fact, we speak of randomness, random events, probabilities, ... We formalize these concepts with a mathematical theory built upon a series of axioms.

# Events and probabilities

An experiment is equivalent to a series of events that lead to some *outcome*.

An experiment is said to be *random* if it is impossible to predict its outcome with certainty.

The set of possible results is called *sample space* and is denoted by $\Omega$.

An element of $\Omega$ is called *outcome* and is denoted by $\omega$.

# Examples

Coin flipping: $\Omega = \{\omega_1, \omega_2\} = \{Head, Tail\}$

or possibly:$\Omega = \{\omega_1, \omega_2, \omega_3\} = \{Head, Tail, Side\}$

Dice roll:$\Omega = \{\omega_1, \ldots, \omega_6\} = \{1, \ldots, 6\}$

The choice of $\Omega$ is, to some extent, arbitrary as it depends on the ex ante idea of the possible outcomes. This kind of assumptions appears in all mathematical models however.

# Random event

An event is said to be *random* if we know whether or not the event happened once the experiment is done.

A random event $A$ can be identified with the subset of $\Omega$ whose elements induce $A$.

We could thus consider every element of the set $\Pi(\Omega)$ of the parts of $\Omega$ as a random event. However, $\Pi(\Omega)$ is often too large (in particular if $\Omega = \mathbb{R}$), and we prefer a simpler class of the parts of $\Omega$ strictly included in $\Pi(\Omega)$.

We impose stability conditions to this class so that usual (set) operations stay within the class.

# Boolean algebra

The events class, denoted by $\mathcal{A}$, is a Boolean algebra of the parts of $\Omega$, that is, a class of the parts of $\Omega$, including $\Omega$, stable under intersection, union and complement operators.

1. If $A, B \in A$ then $A \cup B \in \mathcal{A}$ and $A \cap B \in \mathcal{A}$.
2. If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$.
3. The empty set $\varnothing \in \mathcal{A}$.

# Reminder: terminology

- "$A$ and $B$" $\Leftrightarrow A \cap B$
- "$A$ or $B$" $\Leftrightarrow A \cup B$
- "not $A$"$\Leftrightarrow A^c = \Omega - A$
- "$A$ and $B$ disjoint" $\Leftrightarrow A \cap B = \varnothing$
- "$A$ or $B$ (with $A, B$ disjoints)" $\Leftrightarrow A + B$
- "$A$ implies $B$"$\Leftrightarrow A \subset B$
- "$A$ but not $B$"$\Leftrightarrow A - B$
- "sure event"$\Leftrightarrow \Omega$
- "impossible event"$\Leftrightarrow \varnothing$

# Probability

By definition of a random experiment, its result cannot be predicted with certainty. Also, we can generally not know if an event $A$ will be realized. We want to give $A$ a weight that would indicate its chance of occurrence.

By intuition:

1. If $A$ and $B$ are disjoint, the chance that $A$ or $B$ happen is given by the sum of the weights of $A$ and $B$.
2. If $(A_n)_{n\in\mathbb{N}}$ is a series of events such that each of them is implied by the next one, and that their simultaneous occurrence is impossible, then the weight of $A_n$ is zero in the limit.

This leads to the definition of a probability on a sample space with a Boolean algebra of the events.

# Definition: Probability

We call *probability* on $(\Omega, \mathcal{A})$ a function defined on $\mathcal{A}$ with its value in $[0, 1]$ verifying these three axioms:

1. Normalization axiom : $P(\Omega) = 1$
2. Additivity axiom (total probability): $P(A + B) = P(A) + P(B)$ for $A \cap B = \varnothing$
3. Continuity axiom: If $A_n \downarrow \varnothing$ as $n \to \infty$, $P(A_n) \to 0$

Notation:

- $A + B = A \cup B$ with $A \cap B = \varnothing$
- $A_n \downarrow \varnothing$ or $\lim_{n\to\infty} A_n = \varnothing$ means that $\forall n \in \mathbb{N}$, $A_{n+1} \subset A_n$ and $\cap_{n\in\mathbb{N}} A_n = \emptyset$

# $\sigma$-algebra

An algebra is stable under a finite number of set operations but not necessarily for countable infinite operations.

If we impose a stable structure for countable set operations, countable intersections and countable unions, we speak of a $\sigma$-*algebra*.

A class of the parts of $\Omega$, including $\Omega$, stable under countable intersections, countables unions and complement is called a $\sigma$-*algebra* of the parts of $\Omega$.

- Trivial $\sigma$-*algebra*: $\mathcal{A} = \{\varnothing, \Omega\}$
- $\mathcal{A} = \{\varnothing, A, A^c, \Omega\}$
- Borel $\sigma$-*algebra* $B_{\mathbb{R}}$ generated by the open intervals of $\mathbb{R}$.

# Conditional probability

The probability $P(B)$ measures the chance of occurrence of $B$.

Suppose that we have the additional information « event $A$ will occur », $A$ having a probability $P(A) > 0$. In general, the chance of occurrence of $B$ changes. In particular,

- If $A \cap B = \varnothing$, there is no change
- If $A \subset B$, $B$ becomes certain.

In the case where $A$ occurs, it is natural to measure the chance of occurrence of some event $B$ with a number proportional to $P(A \cap B)$ , that is, of the form $kP(A \cap B)$, $k > 0$.

For $kP(A \cap \cdot)$ to be a probability, we normalize it using $kP(A \cap \Omega) = kP(A) = 1$, or $P(A)k = 1$.

# Conditional probability (cont'd)

Let a probability space $(\Omega, \mathcal{A}, P)$ and an event $A \subset \mathcal{A} : P(A) > 0$. We call conditional probability of an event $B$ given $A$ the positive number $P(A|B) = \frac{P(A \cap B)}{P(A)}$.

Two events$A$ et $B$ are said to be (stochastically) independent if $P(A \cap B) = P(A)P(B)$. In particular, if one of the event has probability 1 or 0, both events are independent. If $P(A) > 0$, $P(B) > 0$, independence yields $P(B|A) = P(B)$ or $P(A|B) = P(A)$, i.e., knowing that one of the events is realized has no impact on the probability of the other.

# Conditional probability (cont'd)

In general, a family $\{A_i; i \in I\}$ is said to be independent if $P(\bigcap_{i \in J} A_i) = \prod_{i \in J} P(A_i)$ for any finite set $J \subset I$.

Remarks:

- independence of events two by two is not a sufficient condition for independence, but independence induces two by two independence.
- if one has $n$ events ($n > 2$), the condition $P(A_1 \cap \cdots \cap A_n) = P(A_1) \ldots P(A_n)$ is not sufficient for independence.

# Random variable and distribution

We are often interested in the consequences of the result of a random experiment, that is, in functions of the result.

Let $(\Omega, \mathcal{A}, P)$ be a probability space, a random variable is a mapping $X : \Omega \to \mathbb{R}$ such that $\{\omega \in \Omega : X(\omega) \leq x\} \subset \mathcal{A}$ for any $x \in \mathbb{R}$. Such a function is said to be $\mathcal{A}$-*measurable*.

The distribution function (or CDF) of a rv is the function $F : \mathbb{R} \to [0, 1]$ giving the probability that $X$ is no larger than $x$, that is:$F(x) = P(A(x))$ with $A(x) = \{\omega \in \Omega : X(\omega) \leq x\}$. We use the following notation::

$F(x) = P(X \leq x),\ x \in \mathbb{R}.$

# Distribution, probability mass and density

A distribution function $F$ has the following properties:

1. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$
2. if $x < y$, then $F(x) \leq F(y)$
3. $F$ is right continuous, $F(x+h) \to F(x)$ as $h \downarrow 0$.

We consider discrete and contiuous rv separately. A rv is said to be discrete if it only takes its values in a countable subset of $\mathbb{R}$. A discrete rv $X$ has a probability mass $f : \mathbb{R} \to [0,1]$ given by $f(x) = P(X = x)$.

A rv is said to be continuous if its distribution function can be written as $F(x) = \int_{-\infty}^{x} f(u)du$, $x \in \mathbb{R}$, where $f : \mathbb{R} \to [0, \infty)$ is an integrable fonction called probability density function (or pdf).

# Distribution function

Considering two rv $X$ and $Y$, we can consider their joint behavior, that is, the behavior of the random vector $(X, Y)$.

The distribution function of a random vector $X = (X_1, \ldots, X_n)$ is the function $F : \mathbb{R}^n \to [0, 1]$ given by $F(x) = P(X \leq x)$, $x \in \mathbb{R}^n$, or more explicitly by $F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$.

Let us consider the bivariate case $(X, Y)$ in order to simplify notation. We have the following properties:

1. $\lim_{x,y\to-\infty} F(x, y) = 0$, $\lim_{x,y\to\infty} F(x, y) = 1$
2. if $(x_1, y_1) \leq (x_2, y_2)$ then $F(x_1, y_1) \leq F(x_2, y_2)$
3. $F$ is right-continuous, $F(x + u, y + v) \to F(x, y)$ as $u, v \downarrow 0$.

# Distribution function (cont'd)

From the joint distribution functions $F = F_{X,Y}$ of the random vector $(X, Y)$, one can compute the univariate distribution functions $F_X$ et $F_Y$ of $X$ and $Y$ called marginal distribution functions.

$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x, y)$
$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y)$

The rv $X$ and $Y$ are said to be discrete if they only take values in a countable subset of $\mathbb{R}^2$. They have a joint probability mass function $f : \mathbb{R}^2 \to [0, 1]$ given by $f(x, y) = P(X = x, Y = y)$.

The rv $X$ and $Y$ are said to be continuous if their joint distribution function can be written as $F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)dudv$, $x, y \in \mathbb{R}$, with $f : \mathbb{R}^2 \to [0, \infty)$ an integrable function called joint probability density function.

# Definition: discrete random variable

A discrete rv only takes values in a countable subset of $\mathbb{R}$ : $\{x_1, x_2, \dots\}$. Its probability mass function $f : \mathbb{R} \to [0, 1]$ defined by $f(x) = P(X = x)$ satisfies:

1. the set of $x$ such that $f(x) \neq 0$ is countable
2. $\sum_i f(x_i) = 1$, with $x_1, x_2, \dots$ the values such that $f(x) \neq 0$.

The distribution function is a step function whose jumps are located at the possible values of the rv: $F(x) = \sum_{i:x_i \leq x} f(x_i)$.

Two discrete rv are independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for any $x$ and $y$. If $X$ and $Y$ are independent and $g, h : \mathbb{R} \to \mathbb{R}$, then $g(X)$ and $h(Y)$ are also independent.

# Moments of a random variable

The expectation (or mean value, or expected value) of a discrete rv with probability mass function $f$ is defined by
$E[X] = \sum_{x:f(x)>0} xf(x)$.

Similarly, if X has a mass function $f$ and $g : \mathbb{R} \to \mathbb{R}$, then
$E[g(X)] = \sum_{x:f(x)>0} g(x)f(x)$.

If $k$ is a positive integer, the $kth$ moment $\mu_k$ of $X$ is defined by
$\mu_k = E[X^k]$. The $kth$ centered moment $m_k$ of $X$ is defined by
$m_k = E[(X - \mu_1)^k]$.

- Expectation: $\mu = \mu_1 = E[X]$
- Variance: $\sigma^2 = m_2 = E[(X - \mu)^2] = E[X^2] - \mu^2 = \mu_2 - \mu^2$
- Standard deviation: $\sigma = \sqrt{m_2}$

# Properties of expectation

Properties of the expectation operator $E$:

1. If $X \geq 0$, then $E[X] \geq 0$
2. If $a, b \in \mathbb{R}$, then $E[aX + bY] = aE[X] + bE[Y]$ (linearity)
3. The rv $X$ always having the value 1 has expectation $E[X] = 1$
4. If $X$ and $Y$ are independent, then $E[XY] = E[X]E[Y]$

Two variables $X$ and $Y$ are said to be uncorrelated if $E[XY] = E[X]E[Y]$.

# Properties of variance, covariance and correlation

Properties of the variance operator $Var$:

1. $Var[aX] = a^2 Var[X]$, $a \in \mathbb{R}$
2. $Var[X+Y] = Var[X] + Var[Y]$ if $X$ and $Y$ are uncorrelated.

The covariace between $X$ and $Y$ is defined by $Cov[X,Y] = E[XY] - E[X]E[Y]$. The correlation coefficient between $X$ and $Y$ is defined by $\rho[X,Y] = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}}$.

- $Cov[X,X] = Var[X]$
- $X$ and $Y$ are uncorrelated if $Cov[X,Y] = 0$.

# Distribution, mass and conditional expectation

The conditional distribution function of $Y$ given $X = x$, denoted $F_{Y|X}(\cdot|x)$, is defined by $F_{Y|X}(y|x) = P(Y \leq y|X = x)$, for any $x$ such that $P[X = x] > 0$.

The conditional probability mass function of $Y$ given $X = x$, denoted $f_{Y|X}(\cdot|x)$,is defined by $f_{Y|X}(y|x) = P(Y = y|X = x)$, for any $x$ such that $P[X = x] > 0$.

- $f_{Y|X} = \frac{f_{X,Y}}{f_X}$
- $X$ and $Y$ are independent if $f_{Y|X} = f_Y$.

The conditional expectation of $Y$ given $X = x$ is the function $\psi(x) = E[Y|X = x] = \sum_y y f_{Y|X}(y|X = x)$. As it depends on the value of $X$, the conditional expectation of $Y$ given $X$, $\psi(X) = E[Y|X]$, is a random variable. We have $E[\psi(X)] = E[Y]$, which means that $E[Y] = \sum_x E[Y|X = x]P[X = x]$.

# Definition

A continuous rv has a distribution function satisfying $F(x) = \int_{-\infty}^{x} f(u)du$ with $f : \mathbb{R} \to [0, \infty)$ integrable and called probability density function of $X$.

1. $\int_{-\infty}^{\infty} f(u)du = 1$
2. $P[X = x] = 0, \forall x \in \mathbb{R}$ (The amount of possible values for a continuous rv is uncountable; it is therefore so large that the probability of any particular value is zero.)
3. $P[a \leq X \leq b] = \int_a^b f(u)du$ (The numerical value $f(x)$ is not a probability. However, we can think of $f(x)dx$ as a probability: $P[x < X < x + dx] = F(x + dx) - F(x) \cong f(x)dx$)

Two continous rv are independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for any $x$ and $y$. If $X$ and $Y$ are independent and $g, h : \mathbb{R} \to \mathbb{R}$, then $g(X)$ and $h(Y)$ are independent too.

# Moments of a random variable

The expectation (or mean value, or expected value) of a continuous rv with pdf $f$ is defined by $E[X] = \int_{-\infty}^{\infty} uf(u)du$.

Similarly, if X has a mass function $f$ and $g : \mathbb{R} \to \mathbb{R}$, then $E[g(X)] = \int_{-\infty}^{\infty} g(u)f(u)du$.

If $k$ is a positive integer, the $kth$ moment $\mu_k$ of $X$ is defined by $\mu_k = E[X^k] = \int_{-\infty}^{\infty} u^k f(u)du$. The $kth$ centered moment $m_k$ of $X$ is defined by $m_k = E[(X - \mu_1)^k]$.

- Expectation: $\mu = \mu_1 = E[X]$
- Variance: $\sigma^2 = m_2 = E[(X - \mu)^2] = E[X^2] - \mu^2 = \mu_2 - \mu^2$
- Standard deviation: $\sigma = \sqrt{m_2}$

# Distribution, density and conditional expectation

The conditional distribution function of $Y$ given $X = x$, denoted $F_{Y|X}(\cdot|x)$, is defined by $F_{Y|X}(y|x) = \int_{-\infty}^{y} \frac{f(x,u)}{f_X(x)} du$, for any $x$ such that $f_X(x) > 0$.

The conditional density function of $Y$ given $X = x$, denoted $f_{Y|X}(\cdot|x)$,is defined by $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$, for all $x$ such that $f_X(x) > 0$.

The conditional expectation of $Y$ given $X = x$ is the function $\psi(x) = E[Y|X = x] = \int_{-\infty}^{\infty} u f_{Y|X}(u|X = x) du$. The random variable $\psi(X) = E[Y|X]$ satisfies in particular $E[\psi(X)] = E[Y]$, which implies that $E[Y] = \int_{-\infty}^{\infty} E[Y|X = u] f_X(u) du$.