Stochastic Models (Lecture #1)

Thomas Verdebout

Université libre de Bruxelles (ULB)

Outline of the course

1. Short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
5. Markov processes, Poisson processes.
6. Brownian motions.

1: Short introduction.

Probability Theory (Part A)

Many phenomena in real life are uncertain.

Simple examples: flip a coin, roll the dice, play roulette, ...

Other examples: weather, Euro-Dollar exchange, time you need to come to the course, how successful is a medication, ...

Any such uncertain phenomenon we shall call experiment.

Probability theory and Mathematical Statistics provide the necessary tools to model and analyze such uncertain phenomena.

Probability Theory (Part A)

Mathematically such an experiment can be described as randomly (whatever that means now) picking an element from a set Ω, say.

Ω (sample space) describes all possible outcomes.

$$\text{Flip a coin} \Longleftrightarrow \Omega = \{H, T\}$$

$$\text{Roll the dice} \Longleftrightarrow \Omega = \{1, 2, 3, 4, 5, 6\}$$

$$\text{Flip a coin 10 times} \Longleftrightarrow \Omega = ?$$

An event is a subset of Ω. $A \subset \Omega$ happens, if we pick ω in A.

The main target: assign to an event A a probability $P(A) \in [0, 1]$.

The mapping $A \mapsto P(A)$ which has to satisfy certain axiomatic properties, is called probability measure.

This looks simple in principle, but it is difficult in practice.

- There can be many events. (e.g. $\Omega = \mathbb{R}$)
- Some logical rules have to hold. (e.g. $A \subset B$ then $P(A) \leq P(B)$)
- $A \mapsto P(A)$ should make sense from a modeling point of view. (e.g. roll the dice $P(\{5\}) = 1/6$).
- ...

Often we are not interested in the outcome ω (the element from Ω which we have picked) itself, but on some other information it contains.

Example: 10 times flipping a coin, how often did we get tail?

A random variable (in the sequel, a *rv*) X can be viewed as a function mapping the set Ω of all possible outcomes ω of a random experiment $\mathcal{E}$ to (real) numbers.

Example: $\omega = (H, H, T, H, T, H, T, T, T, T)$ and $X(\omega) = 6$.

$P(X \leq x)$ is the probability that we pick an ω, such that $X(\omega) \leq x$. Thus $P(X \leq x) = P(\omega | X(\omega) \leq x)$.

Stochastic Processes (Part B)

A stochastic process (in the sequel, a *sp*) $(X_t)_{t \in T}$ is a collection of rv that are all defined on the same random experiment.

Remarks: T is called the index set (usually the “time"). It may be:

- countable ($\rightsquigarrow$ discrete-time sp); e.g., $T = \mathbb{N}$ (ex: price of a stock at the end of day t),
- uncountable ($\rightsquigarrow$ continuous-time sp); e.g., $T = [0, 1]$ (ex: number of clients in a queue at time t).

The state space E is the collection of values the X_t's can assume.

As for the index set, it may be

- countable; e.g., $E = \mathbb{N}$ (ex: number of clients in a queue at time t).
- uncountable; e.g., $E = \mathbb{R}^+$ (ex: water level of a river at distance t from the source).

We recall that for a fixed $t \in T$, X_t is simply a random variable (depending on ω).

For a fixed $\omega \in \Omega$, $t \mapsto X_t(\omega)$ is a function from T into the state space E. This function is called a sample path, a trajectory or a realization of the sp.

$\rightsquigarrow$ It is natural to plot such sample paths . . .

Various sample paths

Annual strikes in the U.S....

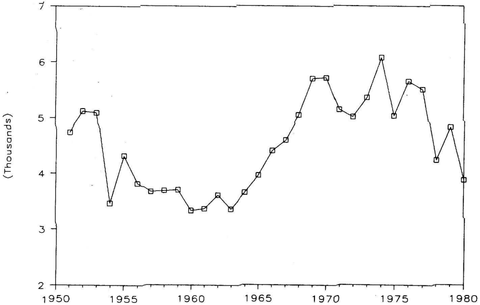
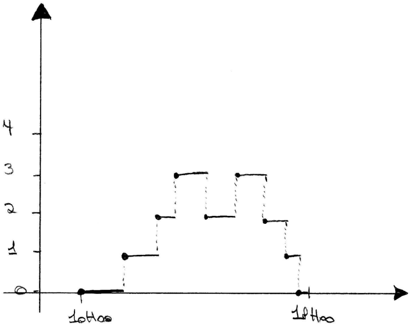


Figure 1.3. Strikes in the U.S.A., 1951–1980 (Bureau of Labor Statistics, U.S. Labor Department).

- Discrete-time process.
- Discrete state space.

Various sample paths

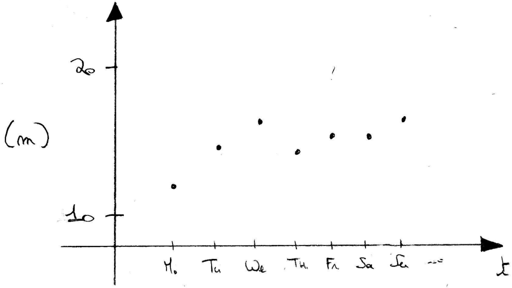
The number of clients in the queue of some shop...



- Continuous-time process.
- Discrete state space.

Various sample paths

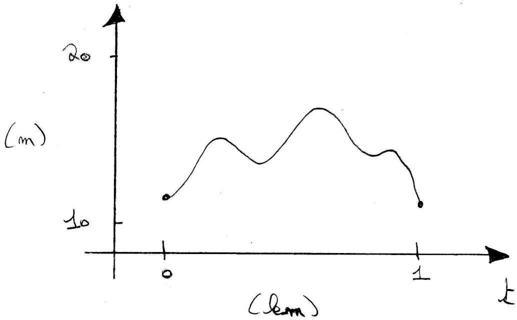
The daily water level (at a fixed location)...



- Discrete-time process.
- Continuous state space.

Various sample paths

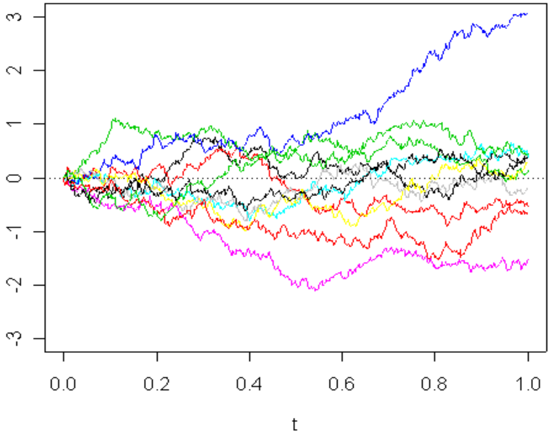
The water level of a river at distance *t* from the source...



- Continuous-time process.
- Continuous state space.

Various sample paths

Trajectories of a standard Brownian motion



- Continuous-time process.
- Continuous state space.

Goals

An observed series $\{x_t, t \in T\}$ can be considered as a realization of some sp $(X_t)_{t \in T}$.

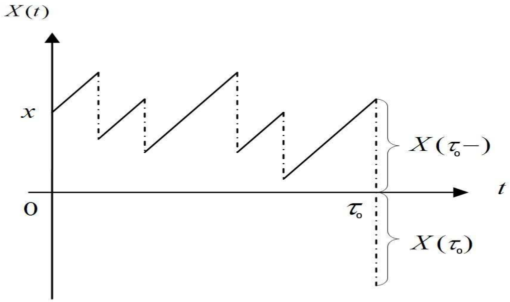
Questions:

- Which assumptions should we put on the X_t's?
- How to model dependence?

Goals:

- Forecasting.
- Strategy.

The ruin problem for an insurance company



where

$$\tau_0 = \inf\{t > 0 : \; X_t < 0\}$$

is a positive-valued random variable which is known as the time of ruin.

Differences w.r.t. statistics and time series courses

Also in statistics one studies random variables and in time series courses one studies stochastic processes.

However, here we assume to know the models which are underlying the experiment.

Statistics and time series is mainly concerned with finding the appropriate model (or parameters) or testing if a model is true.

Illustrative example: flip a coin 10 times.

Probability: what is $P(\#H \geq 8)$, assuming that H and T occur with probability $1/2$ each.

Statistics: we have observed $(H, H, H, T, H, T, H, H, H, H)$. Is the coin fair? What is the probability of H to occur?

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
5. Markov processes, Poisson processes.
6. Brownian motions.

2: Basic probability review.

Random events and probability measure

Consider a random experiment $\mathcal{E}$.

Example: rolling 1 dice, 2 dice, 100m world record attempt, ...

The corresponding sample space Ω is the set of all possible results of $\mathcal{E}$

Example: $\{1, 2, \ldots, 6\}$, $\{(1,1), (1,2), \ldots, (6,6)\}$, $\mathbb{R}_0^+$, ...

An event A can be associated with a subset of Ω.

Example:

- in the 1st $\mathcal{E}$, “obtain 3" $\equiv \{3\}$, “obtain an even result" $\equiv \{2, 4, 6\}$;
- in the 2nd $\mathcal{E}$, “obtain a sum that is ≥ 11" $\equiv \{(5,6), (6,5), (6,6)\}$;
- in the 3rd $\mathcal{E}$, “improving on the world record" $\equiv (0, 9.58)$, ...

Random events and probability measure

We can use set operations to define new events. Let $A_1, A_2, ...$ be events. Then

- $A_1 \cup A_2 \iff A_1$ or A_2 occurs.
- $A_1 \cap A_2 \iff A_1$ and A_2 occurs.
- $A_1 \backslash A_2 \iff A_1$ occurs but A_2 doesn't.
- $A_1^c \iff A_1$ doesn't occur.
- ...
- Generalization to more than two events is obvious.

Random events and probability measure

We want to associate probabilities with events:

$\rightsquigarrow$ A probability measure $\mathbb{P}$ is a function

$$\begin{aligned} \mathbb{P}: \quad \mathcal{A} &\to [0,1] \\ A &\mapsto \mathbb{P}[A], \end{aligned}$$

which satisfies

1. $0 \le \mathbb{P}[A] \le 1$ for all $A \in \mathcal{A}$,
2. $\mathbb{P}[\Omega] = 1$, $\mathbb{P}[\emptyset] = 0$,
3. $\mathbb{P}[\cup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} \mathbb{P}[A_i]$, for all collection of pairwise disjoint $A_1, A_2, \ldots \in \mathcal{A}$, σ-additivity).

Here $\mathcal{A}$ is the collection of all eligible events.

Random events and probability measure

Example: roll a dice.

One can let

- $\mathcal{A} = \mathcal{P}(\Omega)$, i.e., the collection of all subsets of Ω,
- $\mathbb{P}$ defined by $\mathbb{P}[A] = \frac{\#A}{\#\Omega} = \frac{\#A}{6}$.

Remarks:

- One can check that this $\mathbb{P}$ is a probability measure (exercise).
- This choice of $\mathcal{A}$ and $\mathbb{P}$ can be made if $\#\Omega < \infty$.
- Clearly, it cannot be made if $\#\Omega = \infty$. In the general case, we need to restrict the collection of events under consideration ...

  This leads to the (technical) concept of σ-algebra.

σ-algebra

A σ-algebra (or σ-field) over the set Ω is a non-empty collection $\mathcal{A}$ of subsets of Ω satisfying

1. $\Omega \in \mathcal{A}$,
2. $A^c \in \mathcal{A}$, for all $A \in \mathcal{A}$ (where $A^c := \Omega \backslash A$).
3. $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$, for all $A_1, A_2, \ldots \in \mathcal{A}$.

The pair $(\Omega, \mathcal{A})$ is called a measurable space.

The triple $(\Omega, \mathcal{A}, \mathbb{P})$ is called a probability space.

Remarks:

- One can check that a σ-algebra is stable by the standard set operations, such as intersection, union, symmetric difference ...
- A probability measure is always defined on a σ-algebra (which validates the definition of $\mathbb{P}$).

σ-algebra

Usually we want to work with some collection F, say, is of interest. Then one has to work with the σ-algebra $\sigma(F)$ generated by F, i.e. with the smallest σ-algebra containing all elements in F (existence is of course guaranteed).

Most important example:

The so-called Borel σ-field $\mathcal{B}$ over $\Omega = \mathbb{R}$ is

$$\mathcal{B} = \sigma\Big(\big\{(a, b), a < b\big\}\Big).$$

- It can be shown that $\mathcal{B}$ contains all real intervals.
- It clearly plays a central role in probability theory.

Properties of a probability measure

Properties of $\mathbb{P}$ (exercises): Let $A, A_1, A_2 \in \mathcal{A}$, then

- $A_1 \subset A_2$ implies $\mathbb{P}[A_1] \leq \mathbb{P}[A_2]$.
- $\mathbb{P}[A^c] = 1 - \mathbb{P}[A]$.
- $\mathbb{P}[A_1 \cup A_2] = \mathbb{P}[A_1] + \mathbb{P}[A_2]$, if A_1, A_2 are disjoint.
- $\mathbb{P}[A_1 \cup A_2] = \mathbb{P}[A_1] + \mathbb{P}[A_2] - \mathbb{P}[A_1 \cap A_2]$.
- $\mathbb{P}[\cup_{i=1}^{\infty} A_i] \leq \sum_{i=1}^{\infty} \mathbb{P}[A_i]$.
- if $A_1 \subset A_2 \subset ...$, $\mathbb{P}[\cup_{i=1}^{\infty} A_i] = \lim_{i \to \infty} \mathbb{P}[A_i]$.
- if $A_1 \supset A_2 \supset ...$, $\mathbb{P}[\cap_{i=1}^{\infty} A_i] = \lim_{i \to \infty} \mathbb{P}[A_i]$.

Terminology: A occurs a.s. (“almost surely") $\Leftrightarrow \mathbb{P}[A] = 1$.

A very important concept

Often working on the probability space itself is complicated.

We only wish to extract certain information of our experiment.

This leads to the concept of random variables, which basically is nothing but a function $X : \Omega \to \mathbb{R}$.

Random variables

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $(E, \mathcal{B})$ be a measurable space.

A random variable X is a function

$$\begin{aligned} X: \quad \Omega &\to E \\ \omega &\mapsto X(\omega), \end{aligned}$$

which is $\mathcal{A}$-measurable. That is

$$X^{-1}(B) = \{\omega \in \Omega \,|\, X(\omega) \in B\} \in \mathcal{A} \text{ for all } B \in \mathcal{B}. \quad (*)$$

Remark: (*) is the measurability property.

Random variables

Example:

$\mathcal{E}$: flip a coin, $\Omega = \{H, T\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

Some “winning" X may be defined by

$$X(\omega) = \begin{cases} -1 & \text{if} \quad \omega = H \\ 1 & \text{if} \quad \omega = T \end{cases}$$

$\rightsquigarrow$ X is a random variable.

Random variables

Example:

$\mathcal{E}$: 1 dice, $\Omega = \{1, 2, \ldots, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $\mathbb{P}[A] = \frac{\#A}{\#\Omega}$.

Some “winning" X may be defined by

$$X(\omega) = \begin{cases} -30 & \text{if} \quad \omega = 1 \\ -20 & \text{if} \quad \omega = 2 \\ -10 & \text{if} \quad \omega = 3 \\ 10 & \text{if} \quad \omega = 4 \\ 20 & \text{if} \quad \omega = 5 \\ 30 & \text{if} \quad \omega = 6 \end{cases}$$

$\leadsto$ X is a random variable.

Random variables

A fundamental example:

In a general probability space $(\Omega, \mathcal{A}, \mathbb{P})$, the indicator function of the event $A \in \mathcal{A}$ is defined by

$$X(\omega) = \begin{cases} 1 & \text{if} \quad \omega \in A \\ 0 & \text{if} \quad \omega \notin A \end{cases}$$

We usually write $X(\omega) = \mathbb{I}_A(\omega)$.

$\rightsquigarrow$ X is a random variable.

Random variables

Some properties of rv:

- X, Y are rv $\Rightarrow X+Y$, $X-Y$, XY, X/Y, $\max(X, Y)$, and $\min(X, Y)$ are rv
- X is a rv $\Rightarrow X^+ := \max(X, 0)$ and $X^- := -\min(X, 0)$ are rv (and hence $|X| = X^+ + X^-$ is also a rv).
- $X_1, X_2, \ldots$ are rv $\Rightarrow \sup_i X_i$ and $\inf_i X_i$ are rv
- $X_1, X_2, \ldots$ are rv $\Rightarrow \limsup_i X_i$ and $\liminf_i X_i$ are rv
- ...

Definition: The σ-algebra $\sigma(X)$ generated by the rv X is the smallest σ-algebra that makes X measurable (with respect to Borel σ-algebra).

Lemma: $\sigma(X) = X^{-1}(\mathcal{B}) = \{X^{-1}(B) \mid B \in \mathcal{B}\}$.

Distribution

Let X be an $\mathbb{R}$-valued rv defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

The distribution of X is the probability measure $\mathbb{P}_X$ defined on the measurable space $(\mathbb{R}, \mathcal{B})$ by

$$\mathbb{P}_X[B] := \mathbb{P}[X \in B], \quad B \in \mathcal{B}.$$

$\leadsto$ classification: there are “essentially" two types of distributions, discrete and absolutely continuous ones.

Discrete distributions

The distribution $\mathbb{P}_X$ of X is discrete if there exists a (finite or infinite) sequence $x_1, x_2, \ldots \in \mathbb{R}$ such that

$$\mathbb{P}_X[B] = \sum_{i \mid x_i \in B} \mathbb{P}[X = x_i].$$

Such a distribution is characterized by the values x_i

distribution of X			
values	x_1	x_2	$\ldots$
probabilities	p_1	p_2	$\ldots$

Example: consider the rv associated with the winning in the dice game...

Examples of discrete distributions

The Bernoulli distribution
$X \sim \mathrm{Bern}(p)$, with $p \in (0, 1)$:

Bern(p)		
values	0	1
probabilities	$1 - p$	p

Then Binomial distribution
$X \sim \mathrm{Bin}(n, p)$, with $n \in \mathbb{N}_0$ and $p \in (0, 1)$:

Bin(n, p)						
values	0	1	...	k	...	n
probabilities	$(1-p)^n$	$np(1-p)^{n-1}$	...	$\binom{n}{k}p^k(1-p)^{n-k}$	...	p^n

Interpretation: $X \sim \mathrm{Bin}(n, p)$: number of successes in n *independent* similar “$(1, 0)$-trials" (remark: $\mathrm{Bin}(1, p) = \mathrm{Bern}(p)$).

Examples of discrete distributions

The Poisson distribution $X \sim P(\lambda)$, with $\lambda > 0$:

$\mathcal{P}(\lambda)$					
values	0	1	...	k	...
probabilities	$\exp(-\lambda)$	$\exp(-\lambda)\lambda$	...	$\exp(-\lambda)\frac{\lambda^k}{k!}$	...

Interpretation: later...

Absolutely continuous distributions

Let μ and ν be two measures over $(\Omega, \mathcal{A})$. We say that ν is absolutely continuous with respect to μ if for any $A \in \mathcal{A}$, $\mu(A) = 0$ implies $\nu(A) = 0$

The following Theorem plays a central role in the definition of conditional expectations.

Theorem: Let μ and ν be two measures over $(\Omega, \mathcal{A})$ with μ being σ-finite. Then ν is absolutely continuous with respect to μ iff there exists a nonnegative function $f : \Omega \to \mathbb{R}^+$ such that for all $A \in \mathcal{A}$,

$$\nu(A) = \int_A d\nu = \int_A f \, d\mu \tag{1}$$

Absolutely continuous distributions

The function f in the RN Theorem is *essentially unique* in the sense that if f_1 and f_2 are two functions such that (1) hold, then f_1 and f_2 are equal up to μ-measure zero sets.

The set of μ-almost everywhere equal functions is often denoted as $\frac{d\nu}{d\mu}$ and is called the *Radon-Nikodym derivative* of ν with respect to μ.

Absolutely continuous distributions

The distribution $\mathbb{P}_X$ of X is absolutely continuous if, for all $B \in \mathcal{B}$ such that $m[B] = 0$, m being the Lebesgue measure, we have $\mathbb{P}_X[B] = 0$.

It can be shown (Radon-Nikodym theorem) that, for such X, there exists a so-called probability density function (pdf) f such that, for all $B \in \mathcal{B}$,

$$\mathbb{P}_X[B] = \int_B f(x)\,dx.$$

Properties of a pdf: $f(x) \geq 0$ and $\int_{\mathbb{R}} f(x)\,dx = 1$ (of course....)

Such a distribution is characterized by f.

distribution of X	
values	x
pdf	$f(x)$

Examples of absolutely continuous distributions

The Uniform distribution
$X \sim \mathrm{Unif}(a, b)$, with $a, b \in \mathbb{R}$ $(a < b)$:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } x \in (a, b) \\ 0 & \text{otherwise.} \end{cases}$$

The Gaussian (Normal) distribution $X \sim \mathcal{N}(\mu, \sigma^2)$, with $\mu \in \mathbb{R}$ and $\sigma^2 > 0$:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \; x \in \mathbb{R}.$$

Examples of absolutely continuous distributions

The exponential distribution $X \sim \mathrm{Exp}(\lambda)$, with $\lambda > 0$:

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Cumulative distribution

In both cases, one defines the cumulative distribution function (cdf) F of X by

$$F(x) = \mathbb{P}[X \leq x].$$

- In the discrete case, $F(x) = \sum_{i|x_i \leq x} \mathbb{P}[X = x_i] = \sum_{i|x_i \leq x} p_i$, so that F is a step function (exercise).

- In the absolutely continuous case, $F(x) = \int_{-\infty}^{x} f(y)\, dy$, so that F is a continuous function and $F'(x) = f(x)$ (exercise).

In both cases, the cdf satisfies

1. $F(-\infty) = 0$ and $F(\infty) = 1$.
2. F is non-decreasing.
3. F is right-continuous.

Examples

If $X \sim \mathrm{Unif}(a, b)$, with $a, b \in \mathbb{R}$ ($a < b$):

$$F(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } x \in (a, b) \\ 1 & \text{otherwise.} \end{cases}$$

If $X \sim \mathrm{Exp}(\lambda)$,

$$F(x) = \begin{cases} 1 - \exp(-\lambda x) & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$