Università Commerciale
Luigi Bocconi

Lecture 4 – An Introduction to Stochastic Processes

Prof. Massimo Guidolin

Prep Course in Investments

August-September 2016

Plan of the lecture

- Motivation and definitions
- Filtrations and adapted processes
- Conditional probabilities and expectations
- Martingales
- Markov processes

Motivation and definitions

- A stochastic process is a family of random variables (r.v.'s) with a natural ordering
- The representation of the family requires an index and an index set that indicates the ordering, s.t. $\{X_j\}_{j \in \mathcal{J}}$ or $\{X_t\}_{t \in \mathcal{T}}$
  - In applications the index sets often indicate a particular arrangement in space or in time
  - In finance, we specialize to processes that evolve in time
  - In the case of discrete-time processes, the index set $\mathcal{T}$ is countable (usually with equally-spaced elements)
  - In the case of continuous-time processes, $\mathcal{T}$ is a continuum (usually an interval on a time scale)
    - E.g., a continuous record of bid prices for a stock would represent a continuous-time process, while a daily record of settlement prices for commodity futures would be a process in discrete time
  - Regardless of how time is recorded, any r.v. X_t in the family can be either a discrete or a continuous (or a mixed) r.v.

Filtrations and adapted processes

- The modeling of a stochastic process proceeds, as usual, with reference to a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$
  - $(\Omega, \mathcal{F}, \mathbb{P})$ is the basis to assess frequencies with which events occur
    - As the statistics prep course will reveal, Ω is a collection of "outcomes" of some chance experiment and a σ-field, $\mathcal{F}$, contains all those subsets of Ω whose probabilities we will be able to determine
    - Sets in $\mathcal{F}$ are now "events" and we say that event A "occurs" when an outcome ω in set A is obtained.
    - The σ-field $\mathcal{F}$ contains all countable unions and intersections of some basic class of events of interest, plus their complements
    - A probability measure $\mathbb{P}$ is a (finite) mapping from $\mathcal{F}$ into [0, 1]
    - The following fundamental properties of $\mathbb{P}$ are just assumed (these are called axioms): (i) $\mathbb{P}(\Omega) = 1$; (ii) $\mathbb{P}(A) \geq 0$; (iii) $\mathbb{P}(\cup_{j=1}^{\infty} A_j) = \Sigma_{j=1}^{\infty}\mathbb{P}(A_j)$ whenever $A_j \cap A_{j'} = \emptyset \ \forall j, j'$
  - However, a random draw ω from Ω now produces one realization of the entire process, $\{X_t(\omega)\}_{t \in \mathcal{T}}$ over time

Filtrations and adapted processes

- Moreover, as we, the observers of the process, evolve along with it, we acquire new information that includes, at the least, the current and past values of X_t
  - Thinking of $\mathscr{F}$ and the sub-σ-fields of $\mathscr{F}$ as information structures gives a natural way to represent this evolution as a filtration

- Represented as $\mathcal{F} = \{\mathscr{F}_t\}_{t \in \mathscr{T}}$ a filtration is an increasing family of sub- σ-fields of $\mathscr{F}$, meaning that $\mathscr{F}_s \subseteq \mathscr{F}_t$ for $s < t$
  - The process $\{X_t\}$ is then said to be adapted to $\mathcal{F}$, meaning that each X_t and all its history is $\mathcal{F}$-measurable
    - A mapping f from a measurable space $(\Omega, \mathcal{C})$ to a measurable space $(\Psi, \mathcal{D})$ is said to be a measurable mapping if the inverse image under f of each set in $\mathcal{D}$ is a set in $\mathcal{C}$
    - Equivalently, if $f^{-1}(D) \equiv \{\omega : f(\omega) \in D\} \in \mathcal{C} \quad \forall D \in \mathcal{D}$
    - Knowing that f is a measurable mapping means that we can associate with a statement like "f takes a value in $\mathcal{D}$" a probability, a length, or a value of whatever measure it is that has been defined on $(\Omega, \mathcal{C})$

Filtrations and adapted processes

A filtration represents how information evolves as one learns

- For real-valued functions the condition $f^{-1}((-\infty, x]) \in \mathcal{C}$ for each $x \in \mathbb{R}$ (the set of real numbers) is sufficient for measurability

- Measurability captures the absence of prophetic powers
- As an example, consider the experiment of flipping a coin three times in succession, and let Ω comprise the sequences of heads or tails that could be obtained, as

$$\Omega \equiv \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

- Next, take as the general information structure, $\mathcal{F}$, the collection of all 2^8 subsets of these eight elementary outcomes, and define the probability measure $\mathbb{P}$ as

$$\mathbb{P}(A) = \#(A)/\#(\Omega) = \#(A)/8.$$

  - Here "#" is counting measure that enumerates the elementary outcomes in the set, and A is any element of $\mathcal{F}$
- A filtration represents how information evolves as one learns, successively, what happens on each flip

Filtrations and adapted processes

- We can also define an adapted process $\{X_t\}_{t\in\{0,1,2,3\}}$ that represents the total number of heads obtained after t flips
- Initially, before the first flip, we know only that one of the outcomes in Ω will occur and that an outcome not in Ω will not
- Or, the initial information set is the "trivial" field, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and because we start off with $X_0 = 0$, a known constant that does not depend on the outcome of the experiment, X_0 is $\mathcal{F}_0$-measurable
- Now partition Ω into the exclusive sets

$$\underbrace{\{HHH, HHT, HTH, HTT\}}_{H}, \underbrace{\{THH, THT, TTH, TTT\}}_{T}$$

- Once the first flip is made we acquire the information $\mathcal{F}_1 = \{\emptyset, H, T, \Omega\}$ since we then know whether a head or a tail turned up
- Since $\mathcal{F}_0 \subset \mathcal{F}_1$, knowledge we had at the outset has not been lost, and we have added the knowledge of what happened on the 1^{st} flip
- Because we know how many heads have turned up thus far, the random variable X_1 is $\mathcal{F}_1$-measurable

Filtrations and adapted processes

A stochastic process is adapted to a filtration iff the information structure at each t reveals the entire history of the process

- Similarly, $\mathcal{F}_2$ represents the history of the first two flips, the 2^4 events generated by the finer partition

$$\underbrace{\{HHH, HHT\}}_{HH}, \underbrace{\{HTH, HTT\}}_{HT}, \underbrace{\{THH, THT\}}_{TH}, \underbrace{\{TTH, TTT\}}_{TT}$$

- At this point we still know $\mathcal{F}_1$ since $\mathcal{F}_1 :\subset \mathcal{F}_2$, and therefore both X_1 and X_2 are $\mathcal{F}_2$ -measurable
- Finally, making the last flip gives us $\mathcal{F}_3$, the information structure of the 2^8 events generated by the finest possible partition of Ω
- That tells the full story of the three-flip experiment and determines the values of all of X_0, X_1 , X_2, and X_3
- The discrete-time process $\{X_t\}_{t\in\{0,1,2,3\}}$ is then adapted to the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\{0,1,2,3\}}$ in the sense that the information structure at each *t* reveals the entire history of the process

- Can a process ever be not adapted to $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\{0,1,2,3\}}$ and what does that mean exactly?

Filtrations and adapted processes

- In the example $\mathcal{F}_2$ is not the smallest collection of events that determines X_2, meaning that it is not built from the coarsest possible partition that could tell us the total number of heads after two flips

- This minimal σ-field that is generated by X_2, denoted $\sigma(X_2)$, comprises just the 2^3 sets built up from

$$\{TTH, TTT\}, \{THH, THT, HTH, HTT\}, \{HHH, HHT\}$$

- However, X_1 is not measurable w.r.t. $\sigma(X_2)$ and is therefore not a random variable on the probability space $(\emptyset, \Omega, \sigma(X_2), \mathbb{P})$
  - In fact, $\mathcal{F}_2$ is the smallest information set that determines the entire history of X_t through t = 2; that is, $\mathcal{F}_2 = \sigma(X_0, X_1, X_2)$
  - Similarly, $\sigma(X_3)$ is based on a coarser partition than $\mathcal{F}_3$,

$$\{TTT\}, \{TTH, HTT, THT\}, \{HHT, HTH, THH\}, \{HHH\}$$

but $\mathcal{F}_2 = \sigma(X_0, X_1, X_2, X_3)$

- We refer to a probability space endowed with a filtration as a **filtered probability space** and represent it as $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$
  - Filtered spaces lead to properties for conditional probabilities and expectations

Conditional expectations

Under a filtered probability space, the best guess now (time s) of what our expectation will be at, t is our current expectation

- Thinking of the filtration as evolving information, we can interpret $\mathbb{P}(\cdot|\mathcal{F}_t)$ and $E(\cdot|\mathcal{F}_t)$ as showing how assessments of probabilities and expected values evolve over time

- Thus, for an adapted process $\{X_T\}_{t\in[0,T]\}}$ **we can regard $E(\cdot|\mathcal{F}_t)$ itself as a stochastic process** that shows how the expectation of X_T evolves with new information
- Because the "nestedness" of the filtration structure, $\mathcal{F}_s \subseteq \mathcal{F}_t$ for s < t, implies that nothing is forgotten, a joint implication of this feature and the tower property of conditional expectation is

$$E\left[E(X_T|\mathcal{F}_t)|\mathcal{F}_s\right] = E(X_T|\mathcal{F}_s), 0 \leq s \leq t \leq T$$

- The best guess now (time s) of what our expectation will be at t is our current expectation
  - E.g., today's forecast of next week's weather is our best forecast of tomorrow's forecast of next week's weather

Martingales

- Two important classes of stochastic processes are Markov processes and martingales
- An adapted process $\{X_t\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in \mathcal{T}}\ \mathbb{P})$ is a martingale if

  1. $E|X_t| < \infty$ for all $t \geq 0$ and

  2. $E_s X_t = X_s$ for $0 \leq s \leq t$.

  - The integrability requirement, property 1, just ensures that the conditional expectation in 2 always exists
  - The really crucial feature, property 2, will be referred to as the fair-game property; this takes us back to an example in lecture 1
    - Suppose one starts with a known amount X_0 in capital and undertakes a sequence of "fair" bets at times $t \in T = \{0, 1, 2, \ldots\}$, the outcomes of which determine the capital available at t, X_t
- Sub- and super-martingales are, respectively, favorable games $(E_s(X_t) \geq X_s)$ and unfavorable games $(E_s(X_t) \leq X_s)$

Martingales

- For instance, the wealth processes of gambling houses are submartingales, while those of the patrons are super-martingales
  - Since the inequalities are weak, martingales are also sub- and supermartingales
- An important fact that explains much of martingales' central role in modern probability theory is the **martingale convergence theorem**
- If $\{X_t\}$ is a supermartingale with $\sup_t E|X_t| < \infty$, then the limit of X_t as $t \to \infty$ exists a.s. and is finite
  - In short, supermartingales (therefore martingales) whose expected values are bounded converge to something, either a finite constant or a random variable, and neither diverge nor fluctuate indefinitely
  - Uniform integrability of $\{X_t\}$ implies $\sup_t E|X_t| < \infty$
- Let's return now to Markov processes, which are adapted stochastic processes that are memoryless

Markov processes

- This means that predictions of the future depend only on the present state and not on history

- Formally, an adapted process $\{X_t\}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$ is Markov if for any (Borel) set B

$$\mathbb{P}_s(X_t \in B) = \mathbb{P}\left[X_t \in B | \sigma(X_s)\right] = \mathbb{P}(X_t \in B | X_s)$$

  - The probability of any event given the entire information set at $\mathcal{F}_s$ that is, given $\mathcal{F}_t$, is the same as the probability conditional on knowledge of X_s alone
    - The first-order autoregressive process, $X_t = \phi X_{t-1} + u_t$ where X_0 is F0-measurable and the $\{u_t\}$ are i.i.d. with zero mean is a discrete-time Markov process
    - The Wiener process, $\{W_t\}_{t\geq 0}$, also called Brownian motion, is a continuous-time Markov process with independent, normally-distributed increments

- Diffusion processes are Markov processes with continuous sample paths

Readings

- R. DURRETT, *Essentials of stochastic processes*. Springer, 2012.
- S, ROSS, *Stochastic processes*. Vol. 2. New York: John Wiley & Sons, 1996.