Foundations of Financial Economics
Introduction to stochastic processes

Paulo Brito

[1]pbrito@iseg.ulisboa.pt
University of Lisbon

April 30, 2021

Topics for today

- Filtrations

Information set

- The information set is given by

$$(\Omega, \mathcal{F}, \mathcal{P}), \mathbb{F}, \mathbb{P}$$

- where $\mathbb{F}$ is a **filtration**

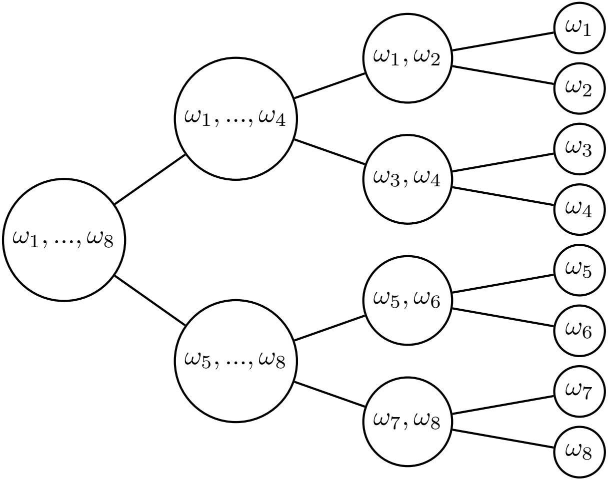
$$\mathbb{F} \equiv \{\mathcal{F}_t\}_{t=0}^{T} = \{\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T\}$$

  which is an ordered sequences of subsets of Ω such that:
  - $\mathcal{F}_0 = \Omega$,
  - $\mathcal{F}_T = \mathcal{F}$ (set of all subsets of Ω)
  - and $\mathcal{F}_t \subset \mathcal{F}_{t+1}$ meaning "more information"
- Then, we can consider a **sequence of events** up until time t

$$W^t = \{\, W_0,\, W_1, \ldots,\, W_t\} \text{ where } W_t \in \mathcal{F}_t$$

Filtration: example

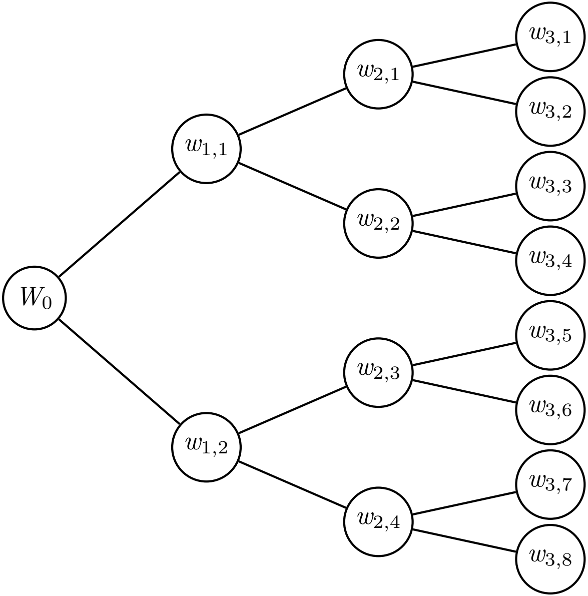
Binomial information tree: for $T = 3$ and $\Omega = \{\omega_1, \ldots, \omega_8\}$



Observation: more information means increasing precision

Filtration: example

Sequence of events: $\{W_0, W_1, W_2, W_3\}$ where $W_1 = \{w_{1,1}, w_{1,2}\}$



Filtration

- at time $t = 0$
  - we observe $W_0 = \Omega$
  - we know that events $w_{1,1}$ or $w_{1,2}$ will occur at time $t = 1$,
  - we also know that
    - if nature picks $w_{1,1}$ events $w_{2,3}$ and $w_{2,4}$, and $w_{3,5}$ to $w_{3,8}$ will not be drawn next
    - if nature picks $w_{1,2}$ events $w_{2,1}$ and $w_{2,2}$, and $w_{3,1}$ to $w_{3,4}$ will not be drawn next
- at time $t = 1$
  - assume that event $w_{1,1}$ has been realized
  - we know that events $w_{2,1}$ or $w_{2,2}$ will occur at time $t = 2$,
  - etc
- this evolution of events are associated to values of random variables and associated probabilities

Stochastic processes

Adapted stochastic processes

- **Definition**: the sequence of random variables X_t

$$X^t = \{X_0, \dots X_t\},\ t \in \mathbb{T}$$

- is called an **adapted stochastic process to the filtration** $\mathbb{F}$ if if X_t is a random variable as regards the event $W_t \in \mathcal{F}_t$, that is

$$X_t = X(W_t),\, W_t \in \mathcal{F}_t$$

- intuition: the information as regards t has the same structure as $\mathcal{F}_t$, in the sense that some potencial sequences are being eliminated across time.

Stochastic processes

Histories

- Let $N^t = \{N_t\}_{t=0}^{T}$, $N_0 = 1$ be the sequence of the number of possible events (which are equal to the number of nodes for an information tree representing $\mathbb{F}$)
- We can represent an adapted stochastic process as a **sequence of possible realizations** for every $t \in 0, \dots, T$

$$X_t = X(W_t) = \begin{pmatrix} x_{t,1} \\ \dots \\ x_{t,N_t} \end{pmatrix} \in \mathbb{R}^{N_t}$$

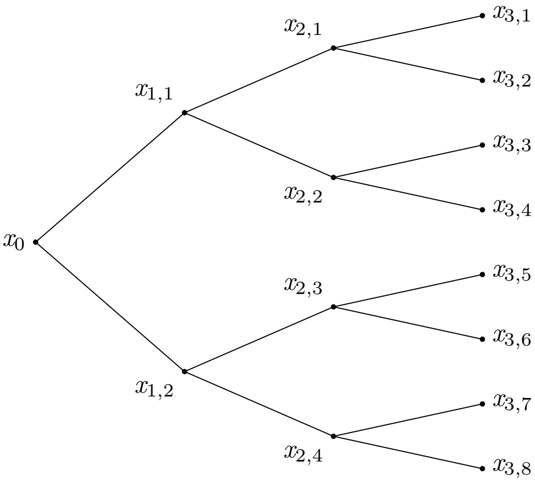
  where N_t is the number of possible realizations of the process at time t,
- **History**: it is a particular realization of $X^t = x^t$ up until time t where

$$x^t = \{x_0, \dots, x_t\}, t \in \mathbb{T}$$

- The set of all histories

$$\mathcal{X}^t = \{X^t\}, \text{ where } X^t = \{X(W_0), X(W_1), \dots X(W_t)\}$$

A binomial stochastic process



- The process $\{X_0, X_1, X_2, X_3\}$ has 8 **possible histories** $\{x_0, x_{1,1}, x_{2,1}, x_{3,1}\}, \ldots \{x_0, x_{1,2}, x_{2,4}, x_{3,8}\}$

Probabilities

- Consider a particular **history** up until time t, $X^t = x^t$

$$x^t = \{x_0, \dots, x_t\}, t \in \mathbb{T}$$

- We call unconditional probability of history x^t to the probability

$$P(x^t) = P(X_t = x_t, X_{t-1}, = x_{t-1}, \dots, X_0 = x_0) \in (0, 1),$$

- Then, we have a sequence of unconditional probability distributions

$$\{\mathsf{P}_0, \mathsf{P}_1, \dots, \mathsf{P}_t\}$$

where $\mathsf{P}_t = P_t(X^t)$ where X^t are **all** histories until time t,

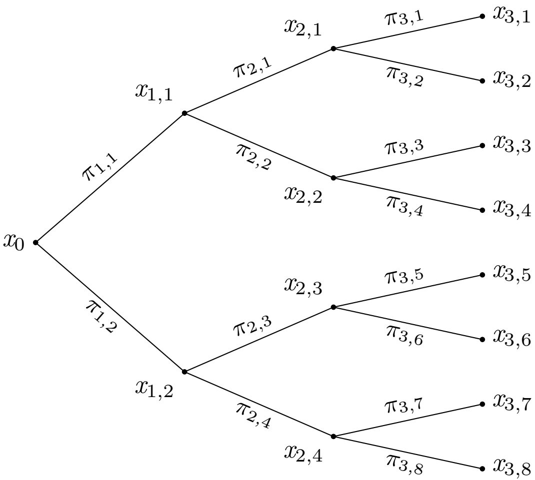
$$\mathsf{P}_t = \begin{pmatrix} \pi_{t,1} \\ \dots \\ \pi_{t,N_t} \end{pmatrix}$$

N_t is the number of nodes of the information at t

- then

$$\mathsf{P}_0 = \sum_{s=1}^{N_t} \pi_{t,s} = 1, \text{for every } t$$

A binomial stochastic process



- The process $\{X_0, X_1, X_2, X_3\}$ has 8 **possible histories**
- The sequence of uncontitional probability distributions is $\{1, \mathsf{P}_1, \mathsf{P}_2, \mathsf{P}_3\}$ where $\sum_{s=1}^{2} \pi_{1,s} = \sum_{s=1}^{4} \pi_{2,s} = \sum_{s=1}^{8} \pi_{3,s} = 1$

Transition probabilities

- The conditional probability of x_{t+1} given a particular history x^t is

$$P(x_{t+1}|x^t) = P(X_{t+1} = x_{t+1}|X_t = x_t, X_{t-1}, = x_{t-1}, \ldots, X_0 = x_0) = \frac{P(X_{t+1} = x_{t+1}, X_t = x_t, X_{t-1}, = x_{t-1}, \ldots, X_0 = x_0)}{P(X_t = x_t, X_{t-1}, = x_{t-1}, \ldots, X_0 = x_0)} \quad (1)$$

- **Definition** we call **transition probability** of $X_{t+h} = x_{t+h}$ given the information history at t,

$$P_t(x_{t+h}) = P(X_{t+h} = x_{t+h}|X^t = x^t)$$

we denote $\mathsf{P}_{t+h|t} = P_t(x_{t+h})$ where

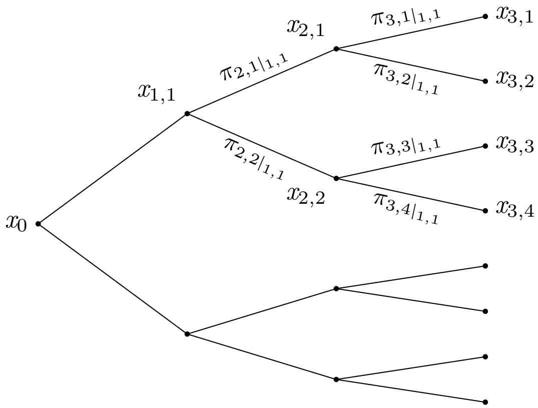
$$\mathsf{P}_{t+h|t} = \begin{pmatrix} \pi_{t+h|t,1} \\ \ldots \\ \pi_{t+h|t,N_{t+h|t}} \end{pmatrix}$$

where $N_{t+h|t}$ is the number of nodes, at $t+h$, of the information node at $x_{t,s}$;

- We have now

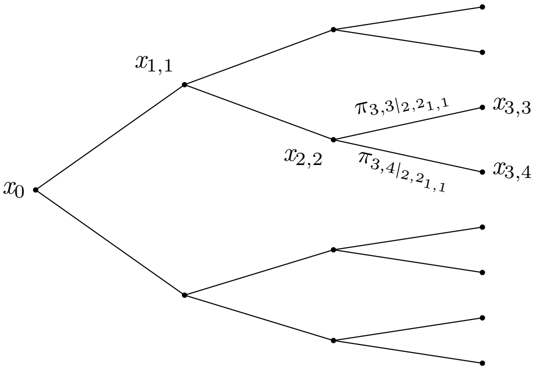
$$\sum_{s=1}^{N_{t+h|t}} \pi_{t+h|t,s} = 1, \text{ for every } t$$

A binomial stochastic process, after a $t = 1$ realization



Conditional probabilities satisfy: $\sum_{s=1}^{2} \pi_{2,s|1,1} = \sum_{s=1}^{4} \pi_{3,s|1,1} = 1$

A binomial stochastic process, after $t = 1$ and $t = 2$ realizations



Conditional probabilities satisfy: $\sum_{s=1}^{2} \pi_{2,s|_{2,2_{1,1}}} = 1$

Markovian processes

- **Definition**: a stochastic process has the **Markov property** if the probability conditional on a **history** is the same as the probability conditional on the **last realization**

$$P(X_{t+h} = x_{t+h} | X^t = x^t) = P(X_{t+h} = x_{t+h} | X_t = x_t)$$

- In other words: the **transition probability** from $X_t = x_t$ is equal to the conditional probability conditional on the history until time t

$$\mathsf{P}_{t+h|t} = P_t(x_{t+h}) \equiv P(X_{t+h} = x_{t+h} | X_t = x_t)$$

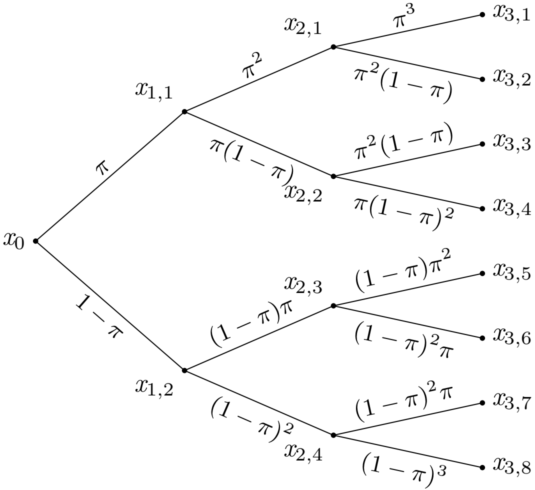
- Observe that a general property of adapted processes is that the unconditional probability of $X_t = x_t$ is equal to the probability of the history x^t, i.e.,

$$\mathsf{P}_t = P_0(x_t) = P(X_t = x_t | X_0 = x_0) = P(x^t)$$

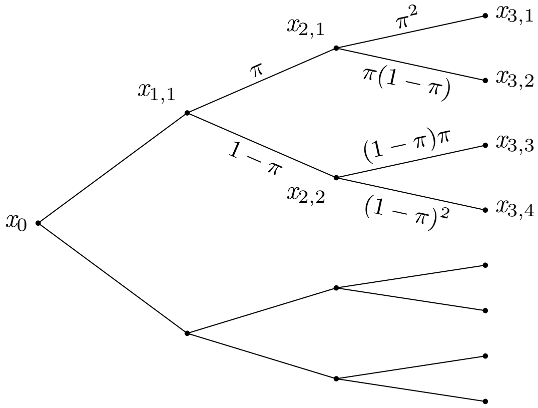
- Then Markov processes verify the following relationship between conditional and unconditional probabilities

$$\mathsf{P}_{t+1} = \mathsf{P}_{t+1|t} \circ \mathsf{P}_t$$

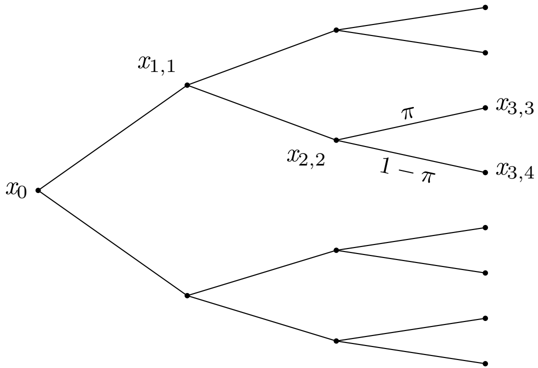
A Markovian binomial process



A Markovian binomial process after a $t = 1$ realization



A Markovian binomial process after a $t = 1$ and $t = 2$ realization



Mathematical expectation for stochastic processes

- **Unconditional mathematical expectation** of X_t is a number

$$\mathbb{E}_0[x_t] = \mathbb{E}[x_t|x_0] = \sum_{s=1}^{N_t} P_0(x_{t,s})x_{t,s} = \sum_{s=1}^{N_t} \pi_{t,s}x_{t,s}$$

- **Unconditional variance** of X_t is

$$\mathbb{V}_0[x_t] = \mathbb{V}[x_t|x_0] = \mathbb{E}_0[(x_t - \mathbb{E}_0(x_t))^2] = \sum_{s=1}^{N_t} \pi_{t,s}(x_{t,s} - \mathbb{E}_0(x_t))^2.$$

- The **conditional mathematical expectation**

$$\mathbb{E}_\tau[x_t] = \mathbb{E}[x_t \mid x^\tau]$$

is an adapted stochastic process because

$$\mathbb{E}_\tau[x_t] = (\mathbb{E}_{\tau,1}(x_t), \dots \mathbb{E}_{\tau,N_\tau}(x_t))$$

where

$$\mathbb{E}_{\tau,i}[x_t] = \sum_{j=1}^{N_{t|\tau,i}} P(X_t = x_{t,j}|x^\tau)x_{t,i} = \sum_{j=1}^{N_{t|\tau,i}} \pi_{t|\tau,j}x_{t,j}, \;\; i = 1, \dots, N_\tau$$

Properties of conditional mathematical expectation: $\mathbb{E}_t$

- if A is a constant
$$\mathbb{E}_t[A] = A$$
- if $X^t = \{x_\tau\}_{\tau=0}^t$ is an adapted process
$$\mathbb{E}_t[x_t] = x_t$$
- **law of the iterated expectations**:
$$\boxed{\mathbb{E}_{t-s}[\mathbb{E}_t[x_{t+r}]] = \mathbb{E}_{t-s}[x_{t+r}],\ s > 0,\ r > 0}$$
this is a very important property: the expected value operator should be taken from the time in which we have the **least** information
- if Y^t is a predictable process (i.e., $\mathcal{F}_{t-1}$-adapted)
$$\mathbb{E}_t[y_{t+1}] = y_{t+1}$$

Martingales

- **Definition**: a process $X^t = \{X_\tau\}_{\tau=0}^t$ has the **martingale property** if

$$\mathbb{E}_t[x_{t+r}] = x_t, \; r > 0$$

- Definition: **super-martingale** if

$$\mathbb{E}_t[x_{t+r}] \leq x_t, \; r > 0$$

- Definition: **sub-martingale** if

$$\mathbb{E}_t[x_{t+r}] \geq x_t, \; r > 0$$

Example

- Let
$$x_{t+1} = \begin{pmatrix} u \times x_t \\ d \times x_t \end{pmatrix} = \begin{pmatrix} u \\ d \end{pmatrix} x_t$$
d and u are known constants such that $0 < d < u$
- and assume that
$$\mathsf{P}_{t+1|t} = \begin{pmatrix} P(x_{t+1} = u \times x_t | x_t) \\ P(x_{t+1} = d \times x_t | x_t) \end{pmatrix} = \begin{pmatrix} p \\ 1-p \end{pmatrix}$$
for $0 < p < 1$
- Then the conditional mathematical expectation is
$$\mathbb{E}_t[x_{t+1}] = (pu + (1-p)d)x_t.$$
- **If** $pu + (1-p)d = 1$ **then** $\mathbb{E}_t[x_{t+1}] = x_t$, that is X^t is a martingale.
- Intuition: the martingale property is associated to the properties of the possible realisations of a stochastic process and of the probability sequence.

Wiener process (or Standard Brownian Motion)

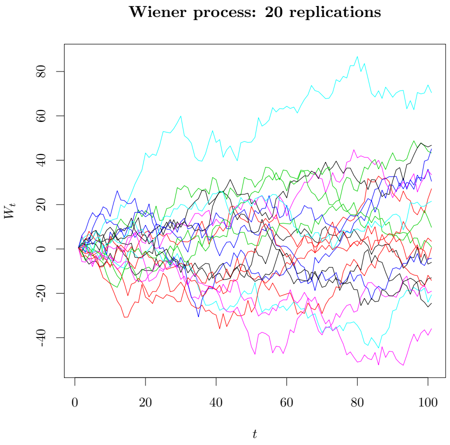
- The process $X^t = \{X_t,\ t \in [0, T)\}$ is a Wiener process if:

$$x_0 = 0,\ \mathbb{E}_0[X_t] = 0,\ \ V_0[X_t - X_\tau] = t - \tau$$

  for any pair $t, \tau \in [0, T)$.
- in particular: $V_0[X_t - X_{t-1}] = 1$
- observe that the process has asymptotically infinite unconditional variance $\lim_{t\to\infty} V_0[X_t - X_\tau] = \infty$ for a finite $\tau \geq 0$
- The variation of the process then follows a stationary standard normal distribution

$$\Delta X_t = X_{t+1} - X_t \sim N(0, 1)$$

Wiener process



Wiener process with drift