Financial derivatives and PDE's
Lecture 3

Simone Calogero

January 21^{th}, 2021

1 Conditional expectation

The **conditional expectation** of a random variable X is an estimate on X based on some information given, for instance, in terms of a sub-σ-algebra $\mathcal{G}$.

Assume first that X is the simple random variable

$$X = \sum_{k=1}^{N} a_k \mathbb{I}_{A_k}.$$

Let $A \in \mathcal{F}$ be an event with $\mathbb{P}(A) > 0$. The conditional expectation of the simple random variable X given the event to A is given by

$$\mathbb{E}[X|A] = \sum_{k=1}^{N} a_k \mathbb{P}(A_k|A).$$

As

$$X\mathbb{I}_A = \sum_{k=1}^{N} a_k \mathbb{I}_{A_k}\mathbb{I}_A = \sum_{k=1}^{N} a_k \mathbb{I}_{A_k \cap A},$$

the identity $\mathbb{E}[X|A] = \frac{\mathbb{E}[X\mathbb{I}_A]}{\mathbb{P}(A)}$ holds. The latter can be used to define the conditional expectation of any random variable given an event:

for any $X : \Omega \to \mathbb{R}$ random variable with finite expectation, the conditional expectation of X given the event to A is defined as

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[X\mathbb{I}_A]}{\mathbb{P}(A)},$$

The conditional expectation of X given a σ-algebra $\mathcal{G} \subset \mathcal{F}$ is denoted by $\mathbb{E}[X|\mathcal{G}]$ and is defined axiomatically by requiring that

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|A] = \mathbb{E}[X|A], \quad \text{for all } A \in \mathcal{G} \text{ such that } \mathbb{P}(A) > 0. \tag{1}$$

It can be shown that there exists a unique (up to null sets) random variable $\mathbb{E}[X|\mathcal{G}]$ that satisfies (1). Moreover it satisfies the properties in the following theorem:

Theorem 1 (Properties of the conditional expectation). *Let $X, Y \in L^1(\Omega)$ and $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. The following properties hold almost surely:*

(i) Linearity: $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X|\mathcal{G}] + \beta\mathbb{E}[Y|\mathcal{G}]$, *for all* $\alpha, \beta \in \mathbb{R}$;

(ii) Monotonicity: If $X \leq Y$ *then* $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$.

(iii) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$;

(iv) If X *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}[X|\mathcal{G}] = X$;

(v) Tower property: If $\mathcal{H} \subset \mathcal{G}$ *is a sub-*σ*-algebra, then* $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$;

(vi) If $\mathcal{G}$ *consists of trivial events only, then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$;

(vii) If X *is independent of* $\mathcal{G}$, *then* $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$;

(viii) Take it out what is known: If X *is* $\mathcal{G}$*-measurable, then* $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$;

(ix) Jensen's inequality: Given $\phi : \mathbb{R} \to \mathbb{R}$ *convex there holds* $\mathbb{E}[\phi(X)|\mathcal{G}] \geq \phi(\mathbb{E}[X|\mathcal{G}])$;

(x) Independence Lemma: If X *is* $\mathcal{G}$*-measurable and* $\mathcal{Y}$ *is independent of* $\mathcal{G}$, *then for any measurable function* $g : \mathbb{R}^2 \to [0, \infty)$, *the function* $f : \mathbb{R} \to [0, \infty)$ *defined by*

$$f(x) = \mathbb{E}[g(x, Y)]$$

is measurable and moreover

$$\mathbb{E}[g(X, Y)|\mathcal{G}] = f(X).$$

Martingale processes

A stochastic process $\{M(t)\}_{t\geq 0}$ is called a **martingale** relative to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ if it is adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$, $M(t) \in L^1(\Omega)$ for all $t \geq 0$, and

$$\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s), \quad \text{for all } 0 \leq s \leq t, \tag{2}$$

for all $t \geq 0$.

Hence a stochastic process is martingale if the information available up to time s does not help to predict whether the stochastic process will raise or fall after time s. If we want to emphasize that the martingale property is satisfied with respect to the probability measure $\mathbb{P}$, we shall say that $\{M(t)\}_{t\geq 0}$ is a $\mathbb{P}$-martingale.

Note that, by (iii) in Theorem 1, the expectation of a martingale is constant, i.e.,

$$\mathbb{E}[M(t)] = \mathbb{E}[M(0)], \quad \text{for all } t \geq 0. \tag{3}$$

Example.

The Brownian motion is a martingale relative to any non-anticipating filtration. In fact

$$\begin{aligned}\mathbb{E}[W(t)|\mathcal{F}(s)] &= \mathbb{E}[(W(t) - W(s))|\mathcal{F}(s)] + \mathbb{E}[W(s)|\mathcal{F}(s)] \\ &= \mathbb{E}[W(t) - W(s)] + W(s) = W(s),\end{aligned}$$

where we used that $W(t) - W(s)$ is independent of $\mathcal{F}(s)$, and so

$$\mathbb{E}[(W(t) - W(s))|\mathcal{F}(s)] = \mathbb{E}[(W(t) - W(s))]$$

and the fact that $W(s)$ is $\mathcal{F}(s)$-measurable and so

$$\mathbb{E}[W(s)|\mathcal{F}(s)] = W(s).$$

Markov processes

A stochastic process $\{X(t)\}_{t\geq 0}$ is called a **Markov process** with respect to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ if it is adapted to $\{\mathcal{F}(t)\}_{t\geq 0}$ and if for every measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $g(X(t)) \in L^1(\Omega)$, for all $t \geq 0$, there exists a measurable function $f_g : [0, \infty) \times [0, \infty) \times \mathbb{R} \to \mathbb{R}$ such that

$$\mathbb{E}[g(X(t))|\mathcal{F}(s)] = f_g(t, s, X(s)), \quad \text{for all } 0 \leq s \leq t. \tag{4}$$

The function $f_g(t, s, \cdot)$ is called the **transition probability** of $\{X(t)\}_{t\geq 0}$ from time s to time t.

If $f_g(t, s, x) = f_g(t - s, 0, x)$, for all $t \geq s$ and $x \in \mathbb{R}$, we say that the Markov process is **time-homogeneous**.

Remark: $f_g(t, t, x) = g(x)$, because $\mathbb{E}[g(X(t)|\mathcal{F}(t)] = g(X(t))$.

The interpretation is the following: for a Markov process, the conditional expectation of $g(X(t))$ at the future time t depends only on the random variable $X(s)$ at time s, and not on the behavior of the process before or after time s.

For a time-homogeneous Markov process the transition between any two different times is equivalent to a transition starting at $s = 0$.

If there exists a measurable function $p : [0, \infty) \times [0, \infty) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that $y \to p(t, s, x, y)$ is integrable for all $(t, s, x) \in [0, \infty) \times [0, \infty) \times \mathbb{R}$ and

$$f_g(t, s, x) = \int_{\mathbb{R}} g(y)p(t, s, x, y)\, dy, \quad \text{for } 0 \leq s < t, \tag{5}$$

holds for all bounded measurable functions g, then we call p the **transition probability density** of the Markov process.

Theorem 2. *Let $\{X(t)\}_{t\geq 0}$ be a Markov process with transition density $p(t,s,x,y)$ relative to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$. Assume $X(s) = x \in \mathbb{R}$ is a deterministic constant and that $\mathcal{F}(s)$ is the trivial σ-algebra. Assume also that, for all $t \geq s$, $X(t)$ admits the density $f_{X(t)}$. Then*

$$f_{X(t)}(y) = p(t,s,x,y).$$

Proof. By definition of density

$$\mathbb{P}(X(t) \leq z) = \int_{-\infty}^{z} f_{X(t)}(y)\, dy,$$

Letting $X(s) = x$ into (4)-(5) we obtain

$$\mathbb{E}[g(X(t))] = \int_{\mathbb{R}} g(y)p(t,s,x,y)\, dy.$$

Choosing $g = \mathbb{I}_{(-\infty,z]}$, we obtain

$$\mathbb{P}(X(t) \leq z) = \int_{-\infty}^{z} p(t,s,x,y)\, dy,$$

for all $z \in \mathbb{R}$, hence $f_{X(t)}(y) = p(t,s,x,y)$. $\square$

Theorem 3. *Let $\{\mathcal{F}(t)\}_{t\geq 0}$ be a non-anticipating filtration for the Brownian motion $\{W(t)\}_{t\geq 0}$. Then $\{W(t)\}_{t\geq 0}$ is a homogeneous Markov process relative to $\{\mathcal{F}(t)\}_{t\geq 0}$ with transition density $p(t,s,x,y) = p_*(t-s,x,y)$, where*

$$p_*(\tau,x,y) = \frac{1}{\sqrt{2\pi\tau}} e^{-\frac{(y-x)^2}{2\tau}}. \tag{6}$$

Proof. The statement holds for $s = t$, with $f_g(t,t,x) = g(x)$. For $s < t$ we write

$$\mathbb{E}[g(W(t))|\mathcal{F}(s)] = \mathbb{E}[g(W(t) - W(s) + W(s))|\mathcal{F}(s)] = \mathbb{E}[\widetilde{g}(W(s), W(t) - W(s))|\mathcal{F}(s)],$$

where $\widetilde{g}(x,y) = g(x+y)$. Since $W(t) - W(s)$ is independent of $\mathcal{F}(s)$ and $W(s)$ is $\mathcal{F}(s)$ measurable, then we can apply Theorem 1(x). Precisely, letting

$$f_g(t,s,x) = \mathbb{E}[\widetilde{g}(x, W(t) - W(s))],$$

we have

$$\mathbb{E}[g(W(t))|\mathcal{F}(s)] = f_g(t,s,W(s)),$$

which proves that the Brownian motion is a Markov process relative to $\{\mathcal{F}(t)\}_{t\geq 0}$. To derive the transition density we use that $Y = W(t) - W(s) \in \mathcal{N}(0, t-s)$, so that

$$\mathbb{E}[g(x+Y)] = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} g(x+y) e^{-\frac{y^2}{2(t-s)}}\, dy = \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} g(y) e^{-\frac{(y-x)^2}{2(t-s)}}\, dy,$$

hence

$$\mathbb{E}[g(W(t))|\mathcal{F}(s)] = \left[\int_{\mathbb{R}} g(y)p_*(t-s,x,y)\,dy\right]_{x=W(s)},$$

where p_* is given by (6). This concludes the proof of the theorem. $\square$

When p is given by (6), the function

$$u(t,x) = \int_{\mathbb{R}} g(y)p_*(t-s,x,y)\,dy \tag{7}$$

solves the **heat equation** with initial datum g at time $t = s$, namely

$$\partial_t u = \frac{1}{2}\partial_x^2 u, \quad u(s,x) = g(x), \quad t > s,\ x \in \mathbb{R}. \tag{8}$$