Lecture 5 · 27.4.2021 · 14:15–16:00 via Zoom

Conditional expectation (continuation)

Elementary definition: given any $B \in \mathscr{F}$ with $\mathbb{P}(B) > 0$ we defined the probability of A conditional on B as

$$\mathbb{P}_B(A) = \mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

However we observed that this definition runs into problem in various situations.

We also observed that given two discrete random variables X, Y then we can define the function

$$h(y) = \sum_x x \mathbb{P}(X = x|Y = y)$$

for all y such that $\mathbb{P}(Y = y) > 0$ and otherwise $h(y) = 8479473894237947$ (for example).

This function is the conditional expectation of X given the event $\{Y = y\}$: it is indicated by

$$\mathbb{E}[X|Y = y] := h(y) = \sum_x x \mathbb{P}(X = x|Y = y) = \sum_x x \mathbb{P}_{\{Y = y\}}(X = x).$$

This function has the following property: for all (bounded and measurable) functions $g: \mathbb{R} \to \mathbb{R}$:

$$\mathbb{E}[X g(Y)] = \mathbb{E}[h(Y) g(Y)] \tag{1}$$

(remember that $g(Y)(\omega) = (g \circ Y)(\omega) = g(Y(\omega))$, automatically one has that the r.v. $g(Y)$ is measurable with respect to $\sigma(Y)$ or simply Y-measurable).

It is not difficult to see that another function h' with the same property must be such that

$$h'(y) = h(y)$$

for all y such that $\mathbb{P}(Y = y) > 0$ (exercise for you).

The property (1) is the key for a general definition of conditional expectation and conditional probability.

Definition 1. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Let $\mathscr{G} \subseteq \mathscr{F}$ a sub-σ-algebra of $\mathscr{F}$ and let X be a (real) integrable random variable (i.e. $\mathbb{E}[|X|] < \infty$). The a conditional expectation of X with respect to the σ-algebra $\mathscr{G}$ (or given $\mathscr{G}$) is a real random variable Z such that*

1. *Z is $\mathscr{G}$-measurable;*
2. *The equation*

$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Z \mathbb{1}_A] \tag{2}$$

holds for all $A \in \mathscr{G}$.

Remark.

- Conditional expectation encodes the intuitive idea of taking the expected value of X given the informations we have at our disposal i.e. the σ-algebra $\mathscr{G}$. Remember that a σ-algebra is a mathematical object which encode the “notion of information”.

- Note that in general we cannot take $Z = X$ because X is not $\mathscr{G}$-mesurable in general. However a conditional expectation of X wrt. $\mathscr{F}$ is X itself.

- If $\mathscr{G} = \{\emptyset, \Omega\}$ (trivial σ-algebra, no information), in this case we can take $Z = \mathbb{E}[X]$.

- It is enough just to have good guess for Z and then check the above definition. There is no general algorithm to compute conditional expectations.

Proposition 2. *If Z is a cond. exp. of an integrable X given $\mathscr{G}$ then we have*

$$\mathbb{E}[|Z|] \leqslant \mathbb{E}[|X|] < \infty.$$

Moreover if Z and Z' are two cond. exp. for X given $\mathscr{G}$ then we must have $Z = Z'$ almost surely.

Proof. In eq. (2) we can take $A = \{\omega \in \Omega : Z(\omega) \geqslant 0\} = Z^{-1}[0, +\infty) \in \mathscr{G}$ since Z is $\mathscr{G}$-measurable. Then

$$\mathbb{E}[Z \mathbb{1}_{Z \geqslant 0}] = \mathbb{E}[Z \mathbb{1}_{Z \geqslant 0}] = \mathbb{E}[X \mathbb{1}_{Z \geqslant 0}]$$

and similarly we have also

$$\mathbb{E}[Z \mathbb{1}_{Z < 0}] = \mathbb{E}[Z \mathbb{1}_{Z < 0}] = \mathbb{E}[X \mathbb{1}_{Z < 0}]$$

so by summing these two equations we get

$$\mathbb{E}[|Z|] = \mathbb{E}[Z \mathbb{1}_{Z \geqslant 0}] - \mathbb{E}[Z \mathbb{1}_{Z < 0}] = \mathbb{E}[X \mathbb{1}_{Z \geqslant 0}] - \mathbb{E}[X \mathbb{1}_{Z < 0}] = \mathbb{E}\Big[X \underbrace{(\mathbb{1}_{Z \geqslant 0} - \mathbb{1}_{Z < 0})}_{\text{sign}(Z)}\Big]$$

and now since $|\text{sign}(Z)| \leqslant 1$ we have

$$|\mathbb{E}[X \,\text{sign}(Z)]| \leqslant \mathbb{E}[|X \,\text{sign}(Z)|] \leqslant \mathbb{E}[|X|].$$

This proves the first statement. For the almost sure uniqueness one observe that if Z, Z' are two conditional expectations for X given $\mathscr{G}$ then we must have

$$\mathbb{E}[Z' \mathbb{1}_A] = \mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Z \mathbb{1}_A]$$

for all $A \in \mathscr{G}$. Therefore $\mathbb{E}[(Z - Z')\mathbb{1}_A] = 0$ for all $A \in \mathscr{G}$ and if we take $A = \{Z - Z' \geqslant 0\} \in \mathscr{G}$ since $Z - Z'$ is $\mathscr{G}$-measurable as a difference of two $\mathscr{G}$-measurable r.v. (exercise). In this case we deduce that $0 = \mathbb{E}[(Z - Z')\mathbb{1}_{\{Z - Z' \geqslant 0\}}]$ and similarly we can deduce that $0 = \mathbb{E}[(Z - Z')\mathbb{1}_{\{Z - Z' < 0\}}]$ and as a consequence we have

$$\mathbb{E}[|Z - Z'|] = \mathbb{E}[(Z - Z')\mathbb{1}_{\{Z - Z' \geqslant 0\}}] - \mathbb{E}[(Z - Z')\mathbb{1}_{\{Z - Z' < 0\}}] = 0.$$

This implies $|Z - Z'| = 0$ $\mathbb{P}$-almost surely. □

Remark. By this Proposition we know that the cond. expectation **if it exists** is unique almost surely. (we encoutered this notion of uniqueness already when discussing the Lebesgue spaces $L^p(\Omega, \mathscr{F}, \mathbb{P})$). Given this uniqueness we denote the (unique) conditional expectation Z of X given $\mathscr{G}$ with the symbol

$$\mathbb{E}[X|\mathscr{G}].$$

Note that is it exists $\mathbb{E}[X|\mathscr{G}]: \Omega \to \mathbb{R}$ is a random variable that is almost surely defined. It is the (almost surely) unique $\mathscr{G}$-measurable r.v. that satisfies the equation

$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathscr{G}]\, \mathbb{1}_A]$$

for all $A \in \mathscr{G}$.

Theorem 3. *For any $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$ and any sub-σ-algebra $\mathscr{G} \subseteq \mathscr{F}$ the conditional expectation $\mathbb{E}[X|\mathscr{G}]$ always exists.*

(We will prove it later on).

Remark 4. By the monotone class theorem it is not difficult to show that if Z is a conditional expectation for X given $\mathscr{G}$ (i.e. $Z = \mathbb{E}[X|\mathscr{G}]$ a.s.) then for every bounded $\mathscr{G}$-measurable r.v. $H: \Omega \to \mathbb{R}$ we have

$$\mathbb{E}[XH] = \mathbb{E}[ZH]. \tag{3}$$

Exercise: prove it. The idea is to go from eq. (2) involving indicator functions, to a similar equation involving simple functions $H \in \mathscr{E}_+$ which are $\mathscr{G}$-measurable and then use stability under monotone limits to apply the monotone class theorem.

Example. Let $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}([0,1])$ and

$$\mathscr{F}_1 = \sigma\left(\{[0,1/2], (1/2,1]\}\right) = \{\emptyset, [0,1/2], (1/2,1], [0,1]\}.$$

Here $\mathscr{F}_1$ encode the information on whether the point ω is at the left or right of $1/2$.

In particular the r.v.

$$X_1 := \mathbb{1}_{[0,1/2]}: \Omega \to \mathbb{R}$$

is $\mathscr{F}_1$ measurable and actually $\sigma(X_1) = \mathscr{F}_1$.

Let also

$$X_2 := \mathbb{1}_{[0,1/4]} + \mathbb{1}_{(1/2,3/4]}$$

and $\mathscr{F}_2 = \sigma(X_1, X_2)$ (i.e. the smallest σ-algebra which makes both X_1 and X_2 measurable).

Note that X_2 is **not** $\mathscr{F}_1$ measurable: informally, the information in $\mathscr{F}_1$ is not enough to determine the value of X_2. For example: $X_2^{-1}(\{1\}) = [0,1/4] \cup (1/2,3/4] \notin \mathscr{F}_1$.

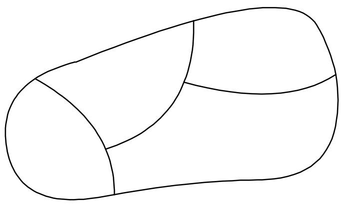
Note that

$$\sigma(X_2) = \{\emptyset, [0,1/4] \cup (1/2,3/4], (1/4,1/2] \cup (3/4,1], \Omega\}.$$

What is $\mathscr{F}_2$? Note that $\mathscr{F}_2 \supseteq \sigma(X_1) \cup \sigma(X_2)$

$$\mathscr{F}_2 = \sigma\left(\{[0,1/4], (1/4,1/2], (1/2,3/4], (3/4,1]\}\right)$$

and indeed $\mathscr{F}_2$ is generated by the partition $\mathscr{P} = \{[0,1/4],(1/4,1/2],(1/2,3/4],(3/4,1]\}$ and it has cardinality $\#\mathscr{F}_2 = \#\sigma(\mathscr{P}) = 2^{\#\mathscr{P}} = 16$.



A r.v. which is measurable wrt. $\mathscr{F}_2$ has to be constant in each of the elements of the partition $\mathscr{P}$.

Let us assume now that $\mathbb{P}$ is the uniform probability on $\Omega = [0,1]$. Then

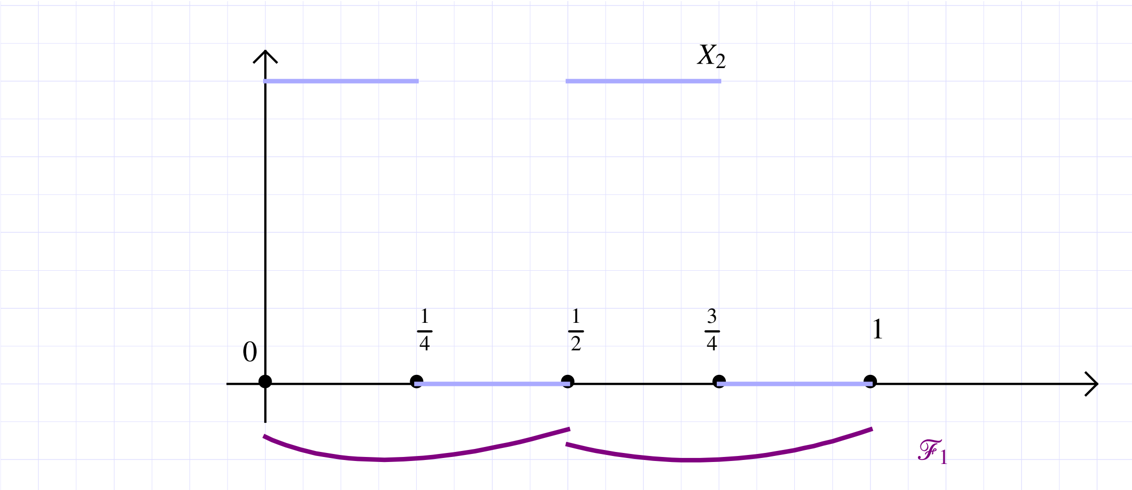
$$\mathbb{P}(X_1 = 1) = \mathbb{P}([0,1/2]) = 1/2, \qquad \mathbb{P}(X_2 = 1) = \mathbb{P}([0,1/4] \cup (1/2,3/4]) = 1/2,$$

so both X_1 and X_2 are Bernoulli random variables of parameter $1/2$. Moreover one can check that they are independent

$$\mathbb{P}(X_1 = x_1, X_2 = x_2) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)$$

for $x_1, x_2 = 0,1$.

Question: what is $\mathbb{E}[X_1|\mathscr{F}_2] = ?$ It is X_1 since X_1 is $\mathscr{F}_2$ measurable. What is $\mathbb{E}[X_2|\mathscr{F}_1] = ??$ Trickier to compute... Intuitively this should be the "average value" of X_2 given that I know $\mathscr{F}_1$ (that is I know if the point ω is $\leqslant 1/2$ or not)



Intuitively we guess that

$$\mathbb{E}[X_2|\mathscr{F}_1] = \frac{1}{2}\mathbb{1}_{[0,1/2]} + \frac{1}{2}\mathbb{1}_{(1/2,1]} = \frac{1}{2}$$

which is indeed the right answer as one can check using the definition.

Let us try now to guess what happens for $\mathbb{E}[X_1X_2|\mathscr{F}_1]$.

$$\mathbb{E}[X_1X_2|\mathscr{F}_1] = ?\mathbb{1}_{[0,1/2]} + ?\mathbb{1}_{(1/2,1]} = \frac{1}{2}\mathbb{1}_{[0,1/2]} + 0\mathbb{1}_{(1/2,1]} = \frac{1}{2}\mathbb{1}_{[0,1/2]} = \frac{1}{2}X_1.$$

On Friday we will prove more properties of cond. exp. in particular we will prove that we can take X_1 out of this expectation, because of measurability and also that conditional expectation of independent r.v. reduce to usual expectation, so that we will be able to show that in this case:

$$\mathbb{E}[X_1 X_2|\mathscr{F}_1] = X_1 \mathbb{E}[X_2|\mathscr{F}_1] = X_1 \mathbb{E}[X_2|\mathscr{F}_1] = X_1 \mathbb{E}[X_2] = \frac{1}{2} X_1,$$

as the specific computation showed.
