Stochastic Processes and Stochastic Calculus - 2
Conditional Expectation

Prof. Maurizio Pratelli

Università degli Studi di Pisa

San Miniato - 12 September 2016

Overview

Back to conditional probability

Recall the definition of conditional probability

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \qquad \text{if } P(B) > 0.$$

Meaning: NEW (updated) probability of A, if the event B occurred.

We can also rewrite independence between A and B as

$$P(A|B) = P(A), \quad \text{if } P(B) > 0.$$

Knowledge the occurrence of B does not affect the degree of plausibility of A.

Back to conditional probability

If we assume to known that B, with $P(B) > 0$, occurred, $P(\cdot|B)$ is a probability measure.

Given a random variable $X : \Omega \to \mathbb{R}$, we have

- the conditional cumulative distribution function of X given B

$$F_X(x \mid B) = P(X \leq x \mid B) = \frac{P(\{X \leq x\} \cap B)}{P(B)}, \qquad x \in \mathbb{R}$$

- the conditional probability distribution of X given B

$$P_{X|B}(A) = P(X \in A \mid B), \qquad A \subseteq \mathbb{R}$$

- the conditional expectation of X given B (if it exists)

$$E[X \mid B] = \frac{1}{P(B)} \int_B X dP = \frac{E[X I_B]}{P(B)}$$

where $I_B : \Omega \to \{0, 1\}$ is the indicator function of B,

$$I_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{if } \omega \notin B \end{cases}$$

Conditional expectation w.r.t. a σ-field – countable case

We generalize $P(\cdot|B)$ from a single event B to a family of events $\mathcal{B}$.

Assume that $\mathcal{B}$ is a σ-field on Ω generated by a countable partition $B_1, B_2, \ldots, B_n, \ldots$

$$\bigcup_{i=1}^{\infty} B_i = \Omega, \quad B_i \cap B_j = \emptyset \quad \text{if } i \neq j.$$

We define, for any (integrable) r.v. $X : \Omega \to \mathbb{R}$, a random variable $E[X \mid \mathcal{B}]$

$$E[X \mid \mathcal{B}](\omega) := E[X|B_i] = \frac{1}{P(B_i)} \int_{B_i} X dP, \quad \text{if } \omega \in B_i.$$

Conditional expectation given a σ-field

$$E[X \mid \mathcal{B}](\omega) := E[X|B_i] = \frac{1}{P(B_i)} \int_{B_i} X dP, \quad \text{if } \omega \in B_i.$$

Some properties

1. $E[E[X \mid \mathcal{B}]] = E[X]$
2. $E[X_1 + X_2 \mid \mathcal{B}] = E[X_1 \mid \mathcal{B}] + E[X_2 \mid \mathcal{B}]$
3. $E[cX \mid \mathcal{B}] = cE[X \mid \mathcal{B}]$ if c is a constant.
4. $E[X \mid \mathcal{B}] = E[X]$ if X and $\mathcal{B}$ are independent.

Conditional expectation w.r.t. a discrete random variable

Suppose now that $Y : \Omega \to E$ is a discrete random variable, i.e. it can take at most a countable number of values $y_1, \dots, y_n, \dots \in E$.
We can let $\mathcal{B} = \sigma(Y)$, which is generated by the countable partition

$$B_1 = \{Y = y_1\} = \{\omega \in \Omega : Y(\omega) = y_1\}$$

$$B_2 = \{Y = y_2\}, \quad B_3 = \{Y = y_3\}, \dots, \quad B_n = \{Y = y_n\}, \dots$$

Given (another) random variable $X : \Omega \to \mathbb{R}$, we define

$$E[X|Y] := E[X|\mathcal{B}]$$

which is a random variable (i.e. defined on Ω).
Notice that on the event $\{Y = y_i\}$, we have

$$E[X|Y](\omega) := E[X|Y = y_i].$$

Therefore we can write

$$E[X|Y] = g(Y),$$

for some $g : E \to \mathbb{R}$.
How to compute g? In the discrete case it is simple:

$$g(y_i) = \frac{1}{P(Y = y_i)} \int_{Y=y_i} X dP.$$

Properties of $E[X|Y]$

1. $E[E[X \mid Y]] = E[X]$
2. $E[X_1 + X_2 \mid Y] = E[X_1 \mid Y] + E[X_2 \mid Y]$
3. $E[cX \mid Y] = cE[X \mid Y]$ if c is a constant.
4. $E[X \mid Y] = E[X]$ if X and Y are independent.

We can improve the last two properties.

In addition to $E[cX \mid Y] = cE[X \mid Y]$, if c is a constant, we have

Proposition

$$E[h(Y)X \mid Y] = h(Y)E[X \mid Y]$$

for any bounded function $h : E \to \mathbb{R}$.

Indeed, for any $\omega \in \Omega$, if $Y(\omega) = y \in E$, we have

$$\begin{aligned} E[h(Y)X \mid Y](\omega) &= E[h(Y)X \mid Y = y] \\ &= E[h(y)X \mid Y = y] \\ &= h(y)E[X \mid Y = y] \\ &= h(Y(\omega))E[X \mid Y](\omega). \end{aligned}$$

In addition to $E[X \mid Y] = E[X]$ if X and Y are independent, we have

Proposition

If X and Y are independent, then

$$E[h(X, Y) \mid Y](\omega) = H(Y(\omega)).$$

for any bounded function $h : \mathbb{R} \times E \to \mathbb{R}$, where

$$H(y) = E[h(X, y)]$$

The formula means that we fix $y \in Y$, compute

$$H(y) = E[h(X, y)]$$

and then evaluate taking $y = Y(\omega)$.
To prove it notice that, for any $\omega \in \Omega$, if $Y(\omega) = y \in E$, we have

$$\begin{aligned} E[h(X, Y) \mid Y](\omega) &= E[h(X, Y) \mid Y = y] \\ &= E[h(X, y) \mid Y = y] \\ &= E[h(X, y)] \\ &= H(y) = H(Y(\omega)). \end{aligned}$$

Problem

How to define $E[X|Y]$ when the random variable Y is not discrete, or when $\mathcal{B}$ is not generated by a countable partition.

The starting point is that the property

$$E[h(Y)X \mid Y] = h(Y)E[X \mid Y]$$

for any bounded function $h : E \to \mathbb{R}$ characterizes the conditional expectation $E[X \mid Y]$.

Indeed if $g(Y)$ is another function such that

$$E[h(Y)X \mid Y] = h(Y)g(Y),$$

for any bounded function $h : E \to \mathbb{R}$, we can take $h(Y) = I_{\{Y=y\}}$ to obtain

$$E[X \mid Y = y] = g(y)$$

Conditional expectation w.r.t. a σ-field

Let $\mathcal{B}$ be a σ-field, $X : \Omega \to \mathbb{R}$ be an integrable random variable.

Definition/Theorem

It is defined the random variable $E[X|\mathcal{B}]$ such that

1. it is $\mathcal{B}$ measurable, i.e. $\{\omega \in \Omega : E[X|\mathcal{B}] \leq x\} \in \mathcal{B}$ for $x \in \mathbb{R}$,
2. it satisfies, for every $B \in \mathcal{B}$,

$$\int_B E[X|\mathcal{B}]\, dP = \int_B X dP.$$

When $\mathcal{B} = \sigma(Y)$, we obtain that $E[X|Y]$

1. is a function of Y, i.e. $E[X|Y] = g(Y)$ ($\to$ how to compute g?)
2. it satisfies, for $B = \{Y \in A\}$ with $P(Y \in A) > 0$,

$$\frac{1}{P(Y \in A)} \int_{\{Y \in A\}} E[X|Y]\, dP = E[X|Y \in A].$$

Properties of the conditional expectation with respect to σ-algebras

1. $E[E[X \mid \mathcal{B}]] = E[X]$
2. $E[X_1 + X_2 \mid \mathcal{B}] = E[X_1 \mid \mathcal{B}] + E[X_2 \mid \mathcal{B}]$
3. $E[cX \mid \mathcal{B}] = cE[X \mid \mathcal{B}]$ if c is a constant.
4. $E[X \mid \mathcal{B}] = E[X]$ if X and $\mathcal{B}$ are independent.
5. $E[YX \mid \mathcal{B}] = YE[X \mid \mathcal{B}]$ if Y is bounded and $\mathcal{B}$-measurable
6. $E[h(X, Y) \mid \mathcal{B}] = H(Y)$ if X and $\mathcal{B}$ are independent, Y is $\mathcal{B}$-measurable, where
$$H(y) = E[h(X, y)].$$

Finally, we have

Proposition

If $\mathcal{B} \subseteq \mathcal{H} \subseteq \mathcal{F}$ are σ-algebras

$$E[E[X \mid \mathcal{H}] \mid \mathcal{B}] = E[X \mid \mathcal{B}].$$

How to compute conditional expectations

We know that $E[X|Y] = g(Y)$. A useful notation for $g(y)$ is $E[X|Y = y]$. Beware that $P(Y = y)$ could be zero, so it is NOT

$$E[X|Y = y] = \frac{1}{P(Y = y)} \int_{Y=y} X dP.$$

A formula holds in case (X, Y) have joint density.

Conditional expectation in case of densities

Assume that (X, Y) have joint density $f(x, y)$. Then, for $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ bounded

$$E[h(X, Y)|Y = y] = \frac{1}{f_Y(y)} \int_{\mathbb{R}} h(x, y) f(x, y) dx,$$

where

$$f_Y(y) = \int_{\mathbb{R}} f(x, y) dx.$$

The rule above, together with

$$E[h(X, Y)|Y = y] = E[h(X, y)]$$

when X and Y are independent are enough to compute many cases.

Conditional expectations and information

Recall that, for a set B

$$P(A|B)$$

is the new probability if we know that B occurs.

What is the meaning of $E[X|\mathcal{B}]$?

We think of the σ-algebra $\mathcal{B}$ as information that we might obtain. Then

$$E[X|\mathcal{B}]$$

is the best approximation of X given the information $\mathcal{B}$.
Best in the sense of the quadratic error:

$$\min_Y E\left[(X-Y)^2\right] = E\left[(X - E[X|\mathcal{B}])^2\right],$$

where the minimum runs among all $\mathcal{B}$-measurables Y.
For example, the property

$$\mathcal{B} \subseteq \mathcal{H} \quad \Rightarrow \quad E[E[X \mid \mathcal{H}] \mid \mathcal{B}] = E[X \mid \mathcal{B}]$$

follows from the fact that more information gives better approximations.

Two words on filtrations

We want to introduce a formalization of the natural idea that

a time goes on, the amout of information increases.

We encode information into σ-algebras.

Definition

Given a "time" set

$$\mathcal{T} = [0, T] \quad \text{(continuous time)}, \qquad \mathcal{T} = \{0, 1, \dots, N\} \text{ (discrete time)},$$

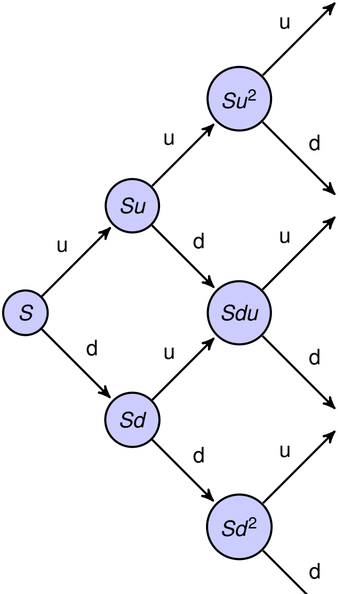
a filtration is a family of σ-fields $(\mathcal{F}_t)_{t \in \mathcal{T}}$ such that

$$s < t \quad \Rightarrow \quad \mathcal{F}_s \subseteq \mathcal{F}_t.$$

An important case is the natural filtration of a process $(X_t)_{t \in \mathcal{T}}$, i.e. we let $\mathcal{F}_t$ be all the information about the history of the process up to time t:

$$\mathcal{F}_t = \sigma(X_s \ : \ s \leq t) \quad \text{(natural filtration)}.$$

An example from the C.R.R. model



Fix $0 < d < 1 < u$.

Price at $t = 0$ is S

Price at $t = 1$ is $S_1 \in \{Su, Sd\}$

Price at $t = 2$ is $S_2 \in \{Su^2, Sdu, Sd^2\}$.

$\Omega = \{u, d\}^N, \mathcal{F} = \mathcal{P}(\Omega)$

Filtration: $\mathcal{F}_0 = \{\emptyset, \Omega\}$

$\mathcal{F}_1 = \{\emptyset, \Omega, (u, \cdot), (d, \cdot), \}$

$\mathcal{F}_2 =$
$\{\emptyset, \Omega, (u, u, \cdot), (u, d, \cdot), (d, u, \cdot), (d, d, \cdot)\}$