MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Fall 2013
Lecture 9 10/2/2013

Conditional expectations, filtration and martingales

Content.

1. Conditional expectations
2. Martingales, sub-martingales and super-martingales

1 Conditional Expectations

1.1 Definition

Recall how we define conditional expectations. Given a random variable X and an event A we define $\mathbb{E}[X|A] = \frac{\mathbb{E}[X1\{A\}]}{\mathbb{P}(A)}$.

Also we can consider conditional expectations with respect to random variables. For simplicity say Y is a simple random variable on Ω taking values $y_1, y_2, \ldots, y_n$ with some probabilities

$$\mathbb{P}(\omega : Y(\omega) = y_i) = p_i.$$

Now we define conditional expectation $\mathbb{E}[X|Y]$ as a random variable which takes value $\mathbb{E}[X|Y = y_i]$ with probability p_i, where $\mathbb{E}[X|Y = y_i]$ should be understood as expectation of X conditioned on the event $\{\omega \in \Omega : Y(\omega) = y_i\}$.

It turns out that one can define conditional expectation with respect to a σ-field. This notion will include both conditioning on events and conditioning on random variables as special cases.

Definition 1. *Given Ω, two σ-fields $\mathcal{G} \subset \mathcal{F}$ on Ω, and a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$. Suppose X is a random variable with respect to $\mathcal{F}$ but not necessarily with respect to $\mathcal{G}$, and suppose X has a finite $\mathbb{L}_1$ norm (that is $\mathbb{E}[|X|] < \infty$). The conditional expectation $\mathbb{E}[X|\mathcal{G}]$ is defined to be a random variable Y which satisfies the following properties:*

(a) Y is measurable with respect to $\mathcal{G}$.

(b) For every $A \in \mathcal{G}$, we have $\mathbb{E}[X1\{A\}] = \mathbb{E}[Y1\{A\}]$.

For simplicity, from now on we write $Z \in \mathcal{F}$ to indicate that Z is measurable with respect to $\mathcal{F}$. Also let $\mathcal{F}(Z)$ denote the smallest σ-field such with respect to which Z is measurable.

Theorem 1. *The conditional expectation* $\mathbb{E}[X|\mathcal{G}]$ *exists and is unique.*

Uniqueness means that if $Y' \in \mathcal{G}$ is any other random variable satisfying conditions (a),(b), then $Y' = Y$ a.s. (with respect to measure $\mathbb{P}$). We will prove this theorem using the notion of Radon-Nikodym derivative, the existence of which we state without a proof below. But before we do this, let us develop some intuition behind this definition.

1.2 Simple properties

- Consider the trivial case when $\mathcal{G} = \{\emptyset, \Omega\}$. We claim that the constant value $c = \mathbb{E}[X]$ is $\mathbb{E}[X|\mathcal{G}]$. Indeed, any constant function is measurable with respect to any σ-field So (a) holds. For (b), we have $\mathbb{E}[X1\{\Omega\}] = \mathbb{E}[X] = c$ and $\mathbb{E}[c1\{\Omega\}] = \mathbb{E}[c] = c$; and $\mathbb{E}[X1\{\emptyset\}] = 0$ and $\mathbb{E}[c1\{\emptyset\}] = 0$.

- As the other extreme, suppose $\mathcal{G} = \mathcal{F}$. Then we claim that $X = \mathbb{E}[X|\mathcal{G}]$. The condition (b) trivially holds. The condition (a) also holds because of the equality between two σ-fields.

- Let us go back to our example of conditional expectation with respect to an event $A \subset \Omega$. Consider the associated σ-fields $\mathcal{G} = \{\emptyset, A, A^c, \Omega\}$ (we established in the first lecture that this is indeed a σ-field). Consider a random variable $Y : \Omega \to \mathbb{R}$ defined as

$$Y(\omega) = \mathbb{E}[X|A] = \frac{\mathbb{E}[X1\{A\}]}{\mathbb{P}(A)} \triangleq c_1$$

for $\omega \in A$ and

$$Y(\omega) = \mathbb{E}[X|A^c] = \frac{\mathbb{E}[X1\{A^c\}]}{\mathbb{P}(A^c)} \triangleq c_2$$

for $\omega \in A^c$. We claim that $Y = \mathbb{E}[X|\mathcal{G}]$. First $Y \in \mathcal{G}$. Indeed, assume for simplicity $c_1 < c_2$. Then $\{\omega : Y(\omega) \leq x\} = \emptyset$ when $x < c_1$, $= A$

for $c_1 \leq x < c_2$, $= \Omega$ when $x \geq c_2$. Thus $Y \in \mathcal{G}$. Then we need to check equality $\mathbb{E}[X1\{B\}] = \mathbb{E}[Y1\{B\}]$ for every $B = \emptyset, A, A^c, \Omega$, which is straightforward to do. For example say $B = A$. Then

$$\mathbb{E}[X1\{A\}] = \mathbb{E}[X|A]\mathbb{P}(A) = c_1\mathbb{P}(A).$$

On the other hand we defined $Y(\omega) = c_1$ for all $\omega \in A$. Thus

$$\mathbb{E}[Y1\{A\}] = c_1\mathbb{E}[1\{A\}] = c_1\mathbb{P}(A).$$

And the equality checks.

- Suppose now $\mathcal{G}$ corresponds to some partition $A_1, \ldots, A_m$ of the sample space Ω. Given $X \in \mathcal{F}$, using a similar analysis, we can check that $Y = \mathbb{E}[X|\mathcal{G}]$ is a random variable which takes values $\mathbb{E}[X|A_j]$ for all $\omega \in A_j$, for $j = 1, 2, \ldots, m$. You will recognize that this is one of our earlier examples where we considered conditioning on a simple random variable Y to get $\mathbb{E}[X|Y]$. In fact this generalizes as follows:

- Given two random variables $X, Y : \Omega \to \mathbb{R}$, suppose both $\in \mathcal{F}$. Let $\mathcal{G} = \mathcal{G}(Y) \subset \mathcal{F}$ be the field generated by Y. We define $\mathbb{E}[X|Y]$ to be $\mathbb{E}[X|\mathcal{G}]$.

1.3 Proof of existence

We now give a proof sketch of Theorem 1.

Proof. Given two probability measures $\mathbb{P}_1, \mathbb{P}_2$ defined on the same $(\Omega, \mathcal{F})$, $\mathbb{P}_2$ is defined to be absolutely continuous with respect to $\mathbb{P}_1$ if for every set $A \in \mathcal{F}$, $\mathbb{P}_1(A) = 0$ implies $\mathbb{P}_2(A) = 0$.

The following theorem is the main technical part for our proof. It involves using the familiar idea of change of measures.

Theorem 2 (**Radon-Nikodym Theorem**). *Suppose $\mathbb{P}_2$ is absolutely continuous with respect to $\mathbb{P}_1$. Then there exists a non-negative random variable $Y : \Omega \to \mathbb{R}_+$ such that for every $A \in \mathcal{F}$*

$$\mathbb{P}_2(A) = \mathbb{E}_{\mathbb{P}_1}[Y1\{A\}].$$

Function Y is called Radon-Nikodym (RN) derivative and sometimes is denoted $d\mathbb{P}_2/d\mathbb{P}_1$.

Problem 1. *Prove that Y is unique up-to measure zero. That is if Y' is also RN derivative, then $Y = Y'$ a.s. w.r.t. $\mathbb{P}_1$ and hence $\mathbb{P}_2$.*

We now use this theorem to establish the existence of conditional expectations. Thus we have $\mathcal{G} \subset \mathcal{F}$, $\mathbb{P}$ is a probability measure on $\mathcal{F}$ and X is measurable with respect to $\mathcal{F}$. We will only consider the case $X \geq 0$ such that $\mathbb{E}[X] < \infty$. We also assume that X is not constant, so that $\mathbb{E}[X] > 0$. Consider a new probability measure $\mathbb{P}_2$ on $\mathcal{G}$ defined as follows:

$$\mathbb{P}_2(A) = \frac{\mathbb{E}_{\mathbb{P}}[X1\{A\}]}{\mathbb{E}_{\mathbb{P}}[X]}, \quad A \in \mathcal{G},$$

where we write $\mathbb{E}_{\mathbb{P}}$ in place of $\mathbb{E}$ to emphasize that the expectation operator is with respect to the original measure $\mathbb{P}$. Check that this is indeed a probability measure on $(\Omega, \mathcal{G})$. Now $\mathbb{P}$ also induced a probability measure on $(\Omega, \mathcal{G})$. We claim that $\mathbb{P}_2$ is absolutely continuous with respect to $\mathbb{P}$. Indeed if $\mathbb{P}(A) = 0$ then the numerator is zero. By the Radon-Nikodym Theorem then there exists Z which is measurable with respect to $\mathcal{G}$ such that for any $A \in \mathcal{G}$

$$\mathbb{P}_2(A) = \mathbb{E}_{\mathbb{P}}[Z1\{A\}].$$

We now take $Y = Z\mathbb{E}_{\mathbb{P}}[X]$. Then Y satisfies the condition (b) of being a conditional expectation, since for every set B

$$\mathbb{E}_{\mathbb{P}}[Y1\{B\}] = \mathbb{E}_{\mathbb{P}}[X]\mathbb{E}_{\mathbb{P}}[Z1\{B\}] = \mathbb{E}_{\mathbb{P}}[X1\{B\}].$$

The second part, corresponding to the uniqueness property is proved similarly to the uniqueness of the RN derivative (Problem 1). □

2 Properties

Here are some additional properties of conditional expectations.

Linearity. $\mathbb{E}[aX + Y|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$.

Monotonicity. If $X_1 \leq X_2$ a.s, then $\mathbb{E}[X_1|\mathcal{G}] \leq \mathbb{E}[X_2|\mathcal{G}]$. Proof idea is similar to the one you need to use for Problem 1.

Independence.

Problem 2. *Suppose X is independent from $\mathcal{G}$. Namely, for every measurable $A \subset \mathbb{R}, B \in \mathcal{G}$ $\mathbb{P}(\{X \in A\} \cap B) = \mathbb{P}(X \in A)\mathbb{P}(B)$. Prove that $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.*

Conditional Jensen's inequality. Let ϕ be a convex function and $\mathbb{E}[|X|], \mathbb{E}[|\phi(X)|] < \infty$. Then $\phi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\phi(X)|\mathcal{G}]$.

Proof. We use the following representation of a convex function, which we do not prove (see Durrett [1]). Let

$$A = \{(a, b) \in \mathbb{Q} : ax + b \leq \phi(x), \forall\, x\}.$$

Then $\phi(x) = \sup\{ax + b : (a, b) \in A\}$.

Now we prove the Jensen's inequality. For any pair of rationals $a, b \in \mathbb{Q}$ satisfying the bound above, we have, by monotonicity that $\mathbb{E}[\phi(X)|\mathcal{G}] \geq a\mathbb{E}[X|\mathcal{G}] + b$, a.s., implying $\mathbb{E}[\phi(X)|\mathcal{G}] \geq \sup\{a\mathbb{E}[X|\mathcal{G}] + b : (a, b) \in A\} = \phi(\mathbb{E}[X|\mathcal{G}])$ a.s. □

Tower property. Suppose $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$. Then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2] = \mathbb{E}[X|\mathcal{G}_1]$ and $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$. That is the smaller field wins.

Proof. By definition $\mathbb{E}[X|\mathcal{G}_1]$ is $\mathcal{G}_1$ measurable. Therefore it is $\mathcal{G}_2$ measurable. Then the first equality follows from the fact $\mathbb{E}[X|\mathcal{G}] = X$, when $X \in \mathcal{G}$, which we established earlier. Now fix any $A \in \mathcal{G}_1$. Denote $\mathbb{E}[X|\mathcal{G}_1]$ by Y_1 and $\mathbb{E}[X|\mathcal{G}_2]$ by Y_2. Then $Y_1 \in \mathcal{G}_1, Y_2 \in \mathcal{G}_2$. Then

$$\mathbb{E}[Y_1 1\{A\}] = \mathbb{E}[X 1\{A\}],$$

simply by the definition of $Y_1 = \mathbb{E}[X|\mathcal{G}_1]$. On the other hand, we also have $A \in \mathcal{G}_2$. Therefore

$$\mathbb{E}[X 1\{A\}] = \mathbb{E}[Y_2 1\{A\}].$$

Combining the two equalities we see that $\mathbb{E}[Y_2 1\{A\}] = \mathbb{E}[Y_1 1\{A\}]$ for every $A \in \mathcal{G}_1$. Therefore, $\mathbb{E}[Y_2|\mathcal{G}_1] = Y_1$, which is the desired result. □

An important special case is when $\mathcal{G}_1$ is a trivial σ-field $\{\emptyset, \Omega\}$. We obtain that for every field $\mathcal{G}$

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

3 Filtration and martingales

3.1 Definition

A family of σ-fields $\{\mathcal{F}_t\}$ is defined to be a filtration if $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ whenever $t_1 \le t_2$. We will consider only two cases when $t \in \mathbb{Z}_+$ or $t \in \mathbb{R}_+$. A stochastic process $\{X_t\}$ is said to be adapted to filtration $\{\mathcal{F}_t\}$ if $X_t \in \mathcal{F}_t$ for every t.

Definition 2. *A stochastic process $\{X_t\}$ adapted to a filtration $\{\mathcal{F}_t\}$ is defined to be a* ***martingale*** *if*

1. $\mathbb{E}[|X_t|] < \infty$ *for all t.*
2. $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$*, for all $s < t$.*

When equality is substituted with $\le$, the process is called ***supermartingale****. When it is substituted with $\ge$, the process is called* ***submartingale****.*

Suppose we have a stochastic process $\{X_t\}$ adapted to filtration $\{\mathcal{F}_t\}$ and suppose for some $s' < s < t$ we have $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ and $\mathbb{E}[X_s|\mathcal{F}_{s'}] = X_{s'}$. Then using Tower property of conditional expectations

$$\mathbb{E}[X_t|\mathcal{F}_{s'}] = \mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_s]|\mathcal{F}_{s'}] = \mathbb{E}[X_s|\mathcal{F}_{s'}] = X_{s'}.$$

This means that when the stochastic process $\{X_n\}$ is discrete time it suffices to check $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ for all n in order to make sure that it is a martingale.

3.2 Simple examples

1. **Random walk**. Let $X_n, n = 1, \ldots$ be an i.i.d. sequence with mean μ and variance $\sigma^2 < \infty$. Let $\mathcal{F}_n$ be the Borel σ-algebra on $\mathbb{R}^n$. Then $S_n - \mu n = \sum_{0 \le k \le n} X_k - \mu n$ is a martingale. Indeed S_n is adapted to $\mathcal{F}_n$, and

$$\begin{aligned}\mathbb{E}[S_{n+1} - (n+1)\mu|\mathcal{F}_n] &= \mathbb{E}[X_{n+1} - \mu + S_n - n\mu|\mathcal{F}_n] \\ &= \mathbb{E}[X_{n+1} - \mu|\mathcal{F}_n] + \mathbb{E}[S_n - n\mu|\mathcal{F}_n] \\ &\stackrel{a}{=} \mathbb{E}[X_{n+1} - \mu] + S_n - n\mu \\ &= S_n - n\mu.\end{aligned}$$

   Here in (a) we used the fact that X_{n+1} is independent from $\mathcal{F}_n$ and $S_n \in \mathcal{F}_n$.

2. **Random walk squared**. Under the same setting, suppose in addition $\mu = 0$. Then $S_n^2 - n\sigma^2$ is a martingale. The proof of this fact is very similar.

4 Additional reading materials

- Durrett [1] Section 4.1, 4.2.

References

[1] R. Durrett, *Probability: theory and examples*, Duxbury Press, second edition, 1996.

15.070J / 6.265J Advanced Stochastic Processes
Fall 2013