STATS310B – Lecture 2

Sourav Chatterjee
Scribed by Michael Howes

01/06/22

Contents

1 Recap

Recall that for an integrable X and a sub σ-algebra $\mathcal{G}$, an integrable random variable Y is $\mathbb{E}(X|\mathcal{G})$ if Y is $\mathcal{G}$–measurable and for all $B \in \mathcal{G}$, $\mathbb{E}(X\mathbf{1}_B) = \mathbb{E}(Y\mathbf{1}_B)$. The random variable Y is said to be the conditional expectation of X given $\mathcal{G}$.

Last time we saw that conditional expectations are unique and they exist if X is square integrable. Today we will prove that conditional expectations exist for all integrable X. We will also derive some properties of conditional expectations.

2 Existence

Note that if X is a simple random variable, then X is bounded and hence square integrable. Thus, the conditional expectation of a simple random variable always exists. Like the unconditional expectation, we will use to prove the existence of conditional expectations for arbitrary integrable random variables.

Lemma 1. *Let X_1 and X_2 be simple random variable and let $a, b \in \mathbb{R}$. Then*

$$\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G}).$$

Proof. Let $Y_i = \mathbb{E}(X_i|\mathcal{G})$ for $i = 1, 2$. The random variable $aY_1 + bY_2$ is $\mathcal{G}$–measurable since Y_1 and Y_2 are both $\mathcal{G}$–measurable. Furthermore, for $B \in \mathcal{G}$,

$$\mathbb{E}[(aY_1 + bY_2)\mathbf{1}_B] = a\mathbb{E}[Y_1\mathbf{1}_B] + b\mathbb{E}[Y_2\mathbf{1}_B] = a\mathbb{E}[X_1\mathbf{1}_B] + b\mathbb{E}[X_2\mathbf{1}_B] = \mathbb{E}[(aX_1 + bX_2)\mathbf{1}_B].$$

Thus, by uniqueness

$$\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G}). \qquad \square$$

Lemma 2. *If X is a non-negative simple random variable, then $\mathbb{E}(X|\mathcal{G})$ is non-negative almost surely.*

Proof. Let $Y = \mathbb{E}(X|\mathcal{G})$ and let $B = \{Y < 0\} \in \mathcal{G}$. Thus,

$$\mathbb{E}[Y\mathbf{1}_B] = \mathbb{E}[X\mathbf{1}_B] \geq 0,$$

since $X\mathbf{1}_B \geq 0$. The random variable $Y\mathbf{1}_B$ is non-positive and thus $Y\mathbf{1}_B = 0$ almost surely. Thus, $Y \geq 0$ almost surely. $\qquad \square$

   Note that $\mathbb{E}(X|\mathcal{G})$ need not be simple even if X is simple. For instance consider $(U, V) \sim$ Uniform$([0, 1]^2)$ and $X = \mathbf{1}_{(U,V) \in D}$ where D is the unit disk and $\mathcal{G} = \sigma(U)$.

Lemma 3. *If X is a $[0, \infty]$-valued random variable. Then $\mathbb{E}(X|\mathcal{G})$ exists and is unique almost surely. More linearity holds in the following sense. If X_1, X_2 are $[0, \infty]$-valued random variables and $a, b \geq 0$, then*

$$\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G}).$$

Proof. Recall that for a $[0, \infty]$-valued random variable X, there exists a sequence of non-negative simple random variables $\{X_n\}_{n \geq 1}$ that increase to X. Fix such a sequence and let $Y_n = \mathbb{E}(X_n|\mathcal{G})$. For any n, $X_n - X_{n-1}$ is a non-negative simple random variable and thus

$$Y_n - Y_{n-1} = \mathbb{E}(X_n - X_{n-1}|\mathcal{G}) \geq 0 \text{ a.s..}$$

Let $A_n = \{\omega : Y_n(\omega) \geq Y_{n-1}(\omega)\}$ and $A = \bigcap_{n=1}^{\infty} A_n$. Since $\mathbb{P}(A_n) = 1$ for all n, we can conclude that $\mathbb{P}(A) = 1$. Furthermore, since each Y_n is $\mathcal{G}$ measurable, the set A is in $\mathcal{G}$. Define

$$Y(\omega) = \begin{cases} \lim_{n \to \infty} Y_n(\omega) & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

Since each Y_n is $\mathcal{G}$ measurable and A is in $\mathcal{G}$, the random variable Y is $\mathcal{G}$ measurable. Furthermore, if $B \in \mathcal{G}$, then $A \cap B \in \mathcal{G}$ and by the monotone convergence theorem,

$$\begin{aligned}
\mathbb{E}[Y\mathbf{1}_B] &= \mathbb{E}[Y\mathbf{1}_{B \cap A}] \\
&= \lim_{n \to \infty} \mathbb{E}[Y_n\mathbf{1}_{B \cap A}] \\
&= \lim_{n \to \infty} \mathbb{E}[X_n\mathbf{1}_{B \cap A}] \\
&= \lim_{n \to \infty} \mathbb{E}[X_n\mathbf{1}_B] \\
&= \mathbb{E}[X\mathbf{1}_B].
\end{aligned}$$

Thus, $Y = \mathbb{E}(X|\mathcal{G})$. Uniqueness and linearity follow from the same argument as before. $\square$

   We are now ready to prove that an arbitrary integrable random variable has a conditional expectation.

Proof. Take any integrable X and any sub-σ-algebra $\mathcal{G}$. Let X^+ and X^- be the positive and negative parts of X. Thus, $|X| = X^+ + X^-$ and so $\mathbb{E}[X^+], \mathbb{E}[X^-] < \infty$. The random variables X^+ and X^- are both $[0, \infty]$-valued. Thus, let $Y_1 = \mathbb{E}[X^+|\mathcal{G}]$ and $Y_2 = \mathbb{E}[X^-|\mathcal{G}]$. Note that,

$$\mathbb{E}[Y_1] = \mathbb{E}[Y_1\mathbf{1}_{\Omega}] = \mathbb{E}[X^+\mathbf{1}_{\Omega}] = \mathbb{E}[X^+] < \infty.$$

Thus, Y_1 is integrable and $Y_1 < \infty$ almost surely. An analogous argument gives that Y_2 is integrable and $Y_2 < \infty$ almost surely. Let $Y = Y_1 - Y_2$ on the set where $|Y_1|, |Y_2| < \infty$ and the $Y = 0$ otherwise. We claim that $Y = \mathbb{E}(X|\mathcal{G})$. First note that Y is integrable and $\mathcal{G}$ measurable since both Y_1 and Y_2 are integrable and $\mathcal{G}$ measurable. Furthermore, if $B \in \mathcal{G}$, then

$$\begin{aligned}
\mathbb{E}[Y\mathbf{1}_B] &= \mathbb{E}[Y_1\mathbf{1}_B] - \mathbb{E}[Y_2\mathbf{1}_B] \\
&= \mathbb{E}[X^+\mathbf{1}_B] - \mathbb{E}[X^-\mathbf{1}_B] \\
&= \mathbb{E}[X\mathbf{1}_B]. \hspace{6cm} \square
\end{aligned}$$

3 Properties of conditional expectation

In the previous section we established the existence of conditional expectations for arbitrary integrable random variables. We will now study some properties of conditional expectations. We will begin by showing that many of the theorems proved for integrals have conditional versions. First we have a conditional version of the triangle inequality.

Proposition 1. *If X is integrable, then $\mathbb{E}[X|\mathcal{G}]$ is integrable and furthermore*

$$|\mathbb{E}(X|\mathcal{G})| \leq \mathbb{E}(|X||\mathcal{G}).$$

Proof. Note that

$$\begin{aligned}
|\mathbb{E}(X|\mathcal{G})| &\leq \left|\mathbb{E}(X^+|\mathcal{G})\right| + \left|\mathbb{E}(X^-|\mathcal{G})\right| \\
&= \mathbb{E}(X^+|\mathcal{G}) + \mathbb{E}(X^-|\mathcal{G}) \\
&= \mathbb{E}(|X||\mathcal{G}).
\end{aligned}$$

Thus,

$$\mathbb{E}[|\mathbb{E}(X|\mathcal{G})|] \leq \mathbb{E}[\mathbb{E}(|X||\mathcal{G})] = \mathbb{E}[|X|]. \qquad \square$$

Linearity also holds for general random variables.

Proposition 2. *If X_1, X_2 are integrable random variables, then for all $a, b \in \mathbb{R}$,*

$$\mathbb{E}(aX_1 + bX_2|\mathcal{G}) = a\mathbb{E}(X_1|\mathcal{G}) + b\mathbb{E}(X_2|\mathcal{G}).$$

Proof. As before. $\qquad \square$

Under independence, the conditional expectation of a random variable is constant.

Proposition 3. *If X is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = X$. In particular if X is a constant, then $\mathbb{E}(X|\mathcal{G}) = X$ for all $\mathcal{G}$.*

Proof. Take any $B \in \mathcal{G}$, then by independence

$$\mathbb{E}[X\mathbf{1}_B] = \mathbb{E}[X]\mathbb{E}[\mathbf{1}_B] = \mathbb{E}[\mathbb{E}[X]\mathbf{1}_B].$$

Thus, $X = \mathbb{E}[X]$ almost surely. $\qquad \square$

On the other extreme, there is the case when X is $\mathcal{G}$-measurable.

Proposition 4. *If X is $\mathcal{G}$-measurable, then $\mathbb{E}(X|\mathcal{G}) = X$.*

Proof. By assumption X is $\mathcal{G}$-measurable and $\mathbb{E}[X\mathbf{1}_B] = \mathbb{E}[X\mathbf{1}_B]$ for all $B \in \mathcal{G}$. $\qquad \square$

The familiar integral convergence theorems all have conditional analogs.

Proposition 5 (Monotone convergence theorem). *Suppose $\{X_n\}$ is an increasing sequence of $[0, \infty]$ values random variables such that $X_n \nearrow X$. Then $\mathbb{E}(X_n|\mathcal{G}) \nearrow \mathbb{E}(X|\mathcal{G})$, almost surely.*

Proof. Let $Y_n = \mathbb{E}(X_n|\mathcal{G})$. Since X_n is increasing, Y_n is increasing almost surely. Thus, define Y to be the almost sure limit of Y_n. Let $B \in \mathcal{G}$, we thus have $X_n\mathbf{1}_B \nearrow X\mathbf{1}_B$ and $Y_n\mathbf{1}_B \nearrow Y\mathbf{1}_B$ almost surely. By the standard monotone convergence theorem, we have

$$\mathbb{E}[X\mathbf{1}_B] = \lim_n \mathbb{E}[X_n\mathbf{1}_B] = \lim_n \mathbb{E}[Y_n\mathbf{1}_B] = \mathbb{E}[Y\mathbf{1}_B].$$

Furthermore, Y is $\mathcal{G}$-measurable since each Y_n is $\mathcal{G}$-measurable. $\qquad \square$

Proposition 6 (Fatou's lemma). *If $\{X_n\}$ is a sequence of non-negative random variables, then*

$$\mathbb{E}(\liminf_n X_n | \mathcal{G}) \le \liminf_n \mathbb{E}(X_n | \mathcal{G}),$$

almost surely.

Proof. For each n, let $Y_n = \inf_{k \ge n} X_k$. The sequence Y_n is increasing and each Y_n is non-negative. Thus, by the monotone convergence theorem

$$\mathbb{E}(\liminf_n X_n | \mathcal{G}) = \mathbb{E}(\lim_n Y_n | \mathcal{G})$$
$$= \lim_n \mathbb{E}(Y_n | \mathcal{G}).$$

Furthermore, for each n we have $X_n \ge Y_n$ and so $\mathbb{E}(Y_n | \mathcal{G}) \le \mathbb{E}(X_n | \mathcal{G})$ almost surely. This implies

$$\lim_n \mathbb{E}(Y_n | \mathcal{G}) \le \liminf_n \mathbb{E}(X_n | \mathcal{G}). \qquad \square$$

We next state and prove the dominated convergence theorem for conditional expectations. For integrals, the conclusion of the dominated convergence theorem is a statement about the convergence of numbers. The conclusion of a conditional dominated convergence theorem would be a statement about convergence of random variables. There are multiple types of convergence of random variables. The conditional dominated convergence theorem guarantees both almost sure convergence and L^1 convergence.

Proposition 7. *Let $\{X_n\}$ be a sequence of random variables converges to X almost surely. Suppose that there exists an integrable random variable Z such that $|X_n| \le Z$ almost surely for all n. Then the following hold*

1. *(Almost sure convergence) $\mathbb{E}(X_n | \mathcal{G}) \to \mathbb{E}(X | G)$ almost surely.*

2. *(Almost sure convergence of the difference) $\mathbb{E}(|X_n - X| | \mathcal{G}) \to 0$ almost surely.*

3. *(L^1 convergence) $\mathbb{E}(|\mathbb{E}(X_n | \mathcal{G}) - \mathbb{E}(X | \mathcal{G})|) \to 0$.*

Proof. We first prove 1. Note that $X_n + Z$ is a non-negative random variable for all n. Thus, by Fatou's lemma we have, almost surely,

$$\mathbb{E}(X | \mathcal{G}) + \mathbb{E}(Z | \mathcal{G}) = \mathbb{E}(X + Z | \mathcal{G})$$
$$= \mathbb{E}(\lim_n (X_n + Z) | \mathcal{G})$$
$$\le \liminf_n \mathbb{E}(X_n + Z | \mathcal{G})$$
$$= \liminf_n \mathbb{E}(X_n | \mathcal{G}) + \mathbb{E}(Z | \mathcal{G}).$$

And so

$$\mathbb{E}(X | \mathcal{G}) \le \liminf_n \mathbb{E}(X_n | \mathcal{G}).$$

Likewise, $Z - X_n$ is non-negative for every n and Fatou's implies that, almost surely,

$$\mathbb{E}(Z - X | \mathcal{G}) \le \liminf_n (\mathbb{E}(Z | \mathcal{G}) - \mathbb{E}(X_n | \mathcal{G})) = \mathbb{E}(Z | \mathcal{G}) - \limsup_n \mathbb{E}(X_n | \mathcal{G}).$$

Thus, almost surely,

$$\limsup_n \mathbb{E}(X_n | \mathcal{G}) \le \mathbb{E}(X | \mathcal{G}) \le \liminf_n \mathbb{E}(X_n | \mathcal{G}),$$

giving 1. Next note that $|X_n - X| \to 0$ almost surely and $|X_n - X| \le 2Z$. Thus, 1. implies 2. Finally, for 3.,

$$\mathbb{E}|\mathbb{E}(X_n|\mathcal{G}) - \mathbb{E}(X|\mathcal{G})| = \mathbb{E}|\mathbb{E}(X_n - X|\mathcal{G})|$$
$$\le \mathbb{E}[\mathbb{E}(|X_n - X||\mathcal{G})]$$
$$= \mathbb{E}[|X_n - X|],$$

and $\mathbb{E}[|X_n - X|] \to 0$ by the unconditional dominated convergence theorem. $\qquad \square$

We now turn to two properties that do not have unconditional analogs. The first is the *tower property*.

Proposition 8. *Suppose X is integrable and $\mathcal{G}' \subseteq \mathcal{G} \subseteq \mathcal{F}$ are both sub-σ-algebra, then*

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{G}') = \mathbb{E}(X|\mathcal{G}').$$

Proof. Let $Y = \mathbb{E}(X|\mathcal{G})$ and $Z = \mathbb{E}(X|\mathcal{G}')$. Let $B \in \mathcal{G}'$ be given and note that $B \in \mathcal{G}$. Thus,

$$\mathbb{E}[Z\mathbf{1}_B] = \mathbb{E}[X\mathbf{1}_B] = \mathbb{E}[Y\mathbf{1}_B].$$

Since Z is $\mathcal{G}'$-measurable, this implies that $Z = \mathbb{E}(Y|\mathcal{G}') = \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{G}')$. $\qquad \square$

The next property is called *taking out what is known*. It shows that $\mathcal{G}$-measurable functions can be treated as constants when calculating conditional expectations with respect to $\mathcal{G}$. The assumptions of the proposition are a little subtle but it is a very powerful result.

Proposition 9. *Suppose that X and Y are two random variables such that*

1. *The random variable Y is $\mathcal{G}$-measurable.*

2. *The random variable X is integrable.*

3. *The product XY is also integrable.*

Then

$$\mathbb{E}(XY|\mathcal{G}) = Y\mathbb{E}(X|\mathcal{G}).$$

Proof. We will fist show that

$$\mathbb{E}[XY] = \mathbb{E}[Y\mathbb{E}(X|\mathcal{G})], \tag{1}$$

for every (X, Y) satisfying the above three conditions. Note that if $Y = \mathbf{1}_B$ for some $B \in \mathcal{G}$, then equation (1) holds by definition of the conditional expectation. By linearity, it follows that equation (1) holds for all $\mathcal{G}$-measurable simple functions. Now suppose that $X, Y \ge 0$ and let Y_n be a sequence of $\mathcal{G}$-measurable, non-negative simple random variables such that $Y_n \nearrow Y$. It follows that $XY_n \nearrow XY$ and $Y_n\mathbb{E}(X_n|\mathcal{G}) \nearrow Y\mathbb{E}(X|\mathcal{G})$. The monotone convergence theorem then gives

$$\mathbb{E}(XY) = \lim_n \mathbb{E}(XY_n) = \lim_n \mathbb{E}(Y_n\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(Y\mathbb{E}(X|\mathcal{G})).$$

Thus, equation (1) is satisfied if X and Y are both non-negative. For general X, Y note that

$$|XY| = (X^+ + X^-)(Y^+ + Y^-) = X^+Y^+ + X^+Y^- + X^-Y^+ + X^-Y^-.$$

Since $\mathbb{E}|XY| < \infty$, each of the random variables X^+Y^+, X^+Y^-, X^-Y^+ and X^-Y^- are also integrable. It follows that

$$\mathbb{E}[XY] = \mathbb{E}[X^+Y^+] - \mathbb{E}[X^+Y^-] - \mathbb{E}[X^-Y^+] + \mathbb{E}[X^-Y^-]$$
$$= \mathbb{E}[\mathbb{E}(X^+|\mathcal{G})Y^+] - \mathbb{E}[\mathbb{E}(X^+|\mathcal{G})Y^-] - \mathbb{E}[\mathbb{E}(X^-|\mathcal{G})Y^+] + \mathbb{E}[\mathbb{E}(X^-|\mathcal{G})Y^-]$$
$$= \mathbb{E}[Y\mathbb{E}(X|\mathcal{G})].$$

Thus, equation (1) holds for arbitrary (X, Y) the above three conditions. We will now see that this implies the theorem. Let X, Y be random variables satisfying 1., 2. and 3. Since Y is $\mathcal{G}$-measurable, $Y\mathbb{E}(X|\mathcal{G})$ is also $\mathcal{G}$-measurable. Let $B \in \mathcal{G}$, the pair $(X, Y\mathbf{1}_B)$ also satisfy the three conditions and so equation (1) also holds. Thus,

$$\mathbb{E}[XY\mathbf{1}_B] = \mathbb{E}[Y\mathbb{E}(X|\mathcal{G})\mathbf{1}_B].$$

Thus, $\mathbb{E}(XY|\mathcal{G}) = Y\mathbb{E}(X|\mathcal{G})$. $\qquad\square$