1 Conditional Expectation

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(X_i)_{i\in I}$ be a collection of random variables defined on this space. Then we define $\sigma(X_i : i \in I) \subseteq \mathcal{F}$ to be the smallest σ-algebra such that all of the X_i are measurable, i.e

$$\sigma(X_i : i \in I) = \sigma(X_i^{-1}(B) : i \in I,\ B \in \mathcal{B}(\mathbb{R})).$$

Definition. If $B \in \mathcal{F}$ has $\mathbb{P}(B) > 0$ then we define

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

for any $A \in \mathcal{F}$. Furthermore, if X is an integrable random variable we define

$$\mathbb{E}[X|B] = \frac{\mathbb{E}[X\mathbb{1}(B)]}{\mathbb{P}(B)}.$$

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say a σ-algebra $\mathcal{G}$ is countably generated if there exist $(B_i)_{i\in I}$ pairwise disjoint (with I countable) such that $\bigcup_{i\in I} B_i = \Omega$ and $\mathcal{G} = \sigma(B_i : i \in I)$.

Let X be an integrable random variable and $\mathcal{G}$ a countably generated σ-algebra. We want to define $X' = \mathbb{E}[X|\mathcal{G}]$. So define

$$X'(\omega) = \mathbb{E}[X|B_i] \text{ whenever } \omega \in B_i.$$

Or equivalently,

$$X'(\omega) = \sum_{i\in I} \mathbb{E}[X|B_i]\mathbb{1}(\omega \in B_i)$$

where we use the convention that $\mathbb{E}[X|B_i] = 0$ if $\mathbb{P}(B_i) = 0$. Then X' is indeed $\mathcal{G}$-measurable (note $\mathcal{G}$ is the set of $\bigcup_{j\in J} B_j$ for $J \subseteq I$).

Note that for any $G \in \mathcal{G}$ we have $\mathbb{E}[X\mathbb{1}(G)] = \mathbb{E}[X'\mathbb{1}(G)]$. Also

$$\mathbb{E}[|X'|] \leq \mathbb{E}\left[\sum_{i\in I} \mathbb{E}[|X||B_i]\mathbb{1}(B_i)\right] = \sum_{i\in I} \mathbb{E}[|X||B_i]\mathbb{P}(B_i) = \mathbb{E}|X| < \infty$$

so X' is integrable.

Theorem (Monotone convergence theorem)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of non-negative random variables with $X_n \uparrow X$ as $n \to \infty$ almost surely. Then $\mathbb{E}X_n \uparrow \mathbb{E}X$ as $n \to \infty$.*

Proof. See Part II Probability & Measure. □

Theorem (Dominated convergence theorem)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of random variables with $X_n \to X$ as $n \to \infty$ almost surely and $|X_n| \leq Y$ almost surely for some Y integrable. Then $\mathbb{E}X_n \to \mathbb{E}X$ as $n \to \infty$.*

Proof. See Part II Probability & Measure. □

Definition (L^p)**.** Let $p \in [1, \infty]$ and f be a measurable function. Define the *L^p-norm*

$$||f||_p = (\mathbb{E}[|f|^p])^{1/p} \text{ for } p \in [1, \infty)$$
$$||f||_\infty = \inf\{\lambda : |f| \leq \lambda \text{ a.e}\}.$$

Furthermore write $f \sim g$ if $f = g$ almost everywhere. Then define the L^p-space $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : ||f||_p < \infty\}/\sim$.

Theorem ($\mathcal{L}^2$ is a Hilbert space)**.** *$\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space with inner product $\langle U, V\rangle = \mathbb{E}[UV]$. For a closed subspace $\mathcal{H}$, if $f \in \mathcal{L}^2$ there exists a unique $g \in \mathcal{H}$ with $||f - g||_2 = \inf\{||f - h||_2 : h \in \mathcal{H}\}$ and $\langle f - g, h\rangle = 0$ for all $h \in \mathcal{H}$. g is called the orthogonal projection of f on $\mathcal{H}$.*

Proof. See Part II Probability & Measure. □

Theorem (Conditional expectation)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-algebra, $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists an integrable random variable Y satisfying*

(a) Y is $\mathcal{G}$-measurable;

(b) for all $A \in \mathcal{G}$, $\mathbb{E}[X\mathbb{1}(A)] = \mathbb{E}[Y\mathbb{1}(A)]$.

Moreover Y is unique, in the sense that if Y' also satisfies (a) and (b), then $Y = Y'$ almost surely. We call Y a version of the conditional expectation of X given $\mathcal{G}$. We write $Y = \mathbb{E}[X|\mathcal{G}]$ almost surely. If $\mathcal{G} = \sigma(Z)$ for a random variable Z, then we write $\mathbb{E}[X|Z] = \mathbb{E}[X|\mathcal{G}]$.

Remark. (b) could be replaced by $\mathbb{E}[XZ] = \mathbb{E}[YZ]$ for all Z bounded and $\mathcal{G}$-measurable.

Proof. First we show uniqueness. Suppose Y and Y' both satisfy (a) and (b) and let $A = \{Y > Y'\} \in \mathcal{G}$. Then

$$\mathbb{E}[Y\mathbb{1}(A)] = \mathbb{E}[Y'\mathbb{1}(A)] \Rightarrow \mathbb{E}[(Y - Y')\mathbb{1}(A)] = 0 \Rightarrow \mathbb{P}(Y > Y') = 0 \Rightarrow Y \leq Y' \text{ a.s}$$

and similarly $Y \geq Y'$ a.s.

Now we show existence. First assume $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Then $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace of $\mathcal{L}^2(\mathcal{F})$. Hence

$$\mathcal{L}^2(\mathcal{F}) = \mathcal{L}^2(\mathcal{G}) \oplus \mathcal{L}^2(\mathcal{G})^\perp$$

so we can write $X = Y + Z$ for $Y \in \mathcal{L}^2(\mathcal{G})$ and $Z \in \mathcal{L}^2(\mathcal{G})^\perp$. Define $\mathbb{E}[X|\mathcal{G}] = Y$, so Y is $\mathcal{G}$-measurable and for all $A \in \mathcal{G}$

$$\mathbb{E}[X\mathbb{1}(A)] = \mathbb{E}[Y\mathbb{1}(A)] + \underbrace{\mathbb{E}[Z\mathbb{1}(A)]}_{=0} = \mathbb{E}[Y\mathbb{1}(A)].$$

We claim that if $X \geq 0$ almost surely, then $Y \geq 0$ almost surely. Indeed, let $A = \{Y < 0\} \in \mathcal{G}$ so $0 \leq \mathbb{E}[X\mathbb{1}(Y < 0)] = \mathbb{E}[Y\mathbb{1}(Y < 0)] \leq 0$ which implies $\mathbb{P}(Y < 0) = 0$.

Assume now that $X \geq 0$ almost surely. Define $X_n = X \wedge n \leq n$, so $X_n \in \mathcal{L}^2$ for all n. Let $Y_n = \mathbb{E}[X_n|\mathcal{G}]$. Then X_n is an increasing sequence and by the above claim, Y_n is also an increasing sequence almost surely. Define $Y = \limsup_{n\to\infty} Y_n$, so Y is $\mathcal{G}$-measurable. Also $Y =\uparrow \lim_{n\to\infty} Y_n$ almost surely. For any $A \in \mathcal{G}$ we have

$$\mathbb{E}[X\mathbb{1}(A)] = \lim_{n\to\infty} \mathbb{E}[X_n\mathbb{1}(A)] = \lim_{n\to\infty} \mathbb{E}[Y_n\mathbb{1}(A)] = \mathbb{E}[Y\mathbb{1}(A)]$$

by the Monotone Convergence Theorem.

Finally, for general X write $X = X^+ - X^-$ and define $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$. □

Remark. From the last proof we can see that we can define $\mathbb{E}[X|\mathcal{G}]$ for $X \geq 0$ without assuming integrability of X. It satisfies all the conditions apart from integrability.

Definition. Let $(\mathcal{G}_n)_{n\geq 1}$ be sub σ-algebras of $\mathcal{F}$. We call them *independent* if whenever $G_i \in \mathcal{G}_i$ and $i_1 < i_2 < \ldots < i_k$ we have

$$\mathbb{P}(G_{i_1} \cap \ldots \cap G_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(G_{i_j}).$$

For a random variable X and a σ-algebra $\mathcal{G}$, we say they are *independent* if $\sigma(X)$ is independent of $\mathcal{G}$.

Properties of conditional expectation

Let $X, Y \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-algebra. Then

1. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$ (take $A = \Omega$);
2. If X is $\mathcal{G}$-measurable then $\mathbb{E}[X|\mathcal{G}] = X$ almost surely (X clearly satisfies the conditions);
3. If X is independent of $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ almost surely;
4. If $X \geq 0$ almost surely then $\mathbb{E}[X|\mathcal{G}] \geq 0$ almost surely;
5. For $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X|\mathcal{G}] + \beta\mathbb{E}[Y|\mathcal{G}]$ almost surely;
6. $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ almost surely.

Recall:

Theorem (Fatou's Lemma)**.** *If $X_n \geq 0$ for all n almost surely, then*

$$\mathbb{E}[\liminf_{n\geq 1} X_n] \leq \liminf_{n\geq 1} \mathbb{E}X_n.$$

Proof. See Part II Probability & Measure. □

Theorem (Jensen's Inequality)**.** *If X is integrable, $\varphi : \mathbb{R} \to \mathbb{R}$ is convex, then*

$$\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X]).$$

We consider any analogues of our convergence theorems for conditional expectation.

Theorem (Conditional Monotone Convergence Theorem)**.** *Suppose $X_n \geq 0$ for all n and $X_n \uparrow X$ almost surely as $n \to \infty$. Let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. Then $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ almost surely.*

Remark. Note that $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ in the almost-sure sense, as these are random variables.

Proof. Let $Y_n = \mathbb{E}[X_n|\mathcal{G}]$ almost surely. Then Y_n is increasing. Set $Y = \limsup_{n\geq 1} Y_n$. Since Y_n is $\mathcal{G}$-measurable, Y is $\mathcal{G}$-measurable. Also $Y =\uparrow \lim_{n\geq 1} Y_n$ almost surely. We need to show $\mathbb{E}[Y\mathbb{1}(A)] = \mathbb{E}[X\mathbb{1}(A)]$ for all $A \in \mathcal{G}$. This follows from the usual Monotone Convergence Theorem as

$$\mathbb{E}[Y\mathbb{1}(A)] = \lim_{n\geq 1} \mathbb{E}[Y_n\mathbb{1}(A)] = \lim_{n\geq 1} \mathbb{E}[X_n\mathbb{1}(A)] = \mathbb{E}[X\mathbb{1}(A)].$$

□

Theorem (Conditional Fatou's Lemma)**.** *Let $(X_n)_{n\geq 1}$ be a non-negative sequence of random variables. Then*

$$\mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \text{ almost surely.}$$

Proof. Note that $\inf_{k\geq n} X_k \uparrow \liminf_{n\to\infty} X_n$ so by the conditional MCT

$$\lim_{n\to\infty} \mathbb{E}[\inf_{k\geq n} X_k|\mathcal{G}] = \mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}].$$

We also have

$$\mathbb{E}[\inf_{k\geq n} X_k|\mathcal{G}] \leq \mathbb{E}[X_k|\mathcal{G}] \quad \forall k \geq n \text{ almost surely.}$$

Which implies

$$\mathbb{E}[\inf_{k\geq n} X_k|\mathcal{G}] \leq \inf_{k\geq n} \mathbb{E}[X_k|\mathcal{G}] \quad \forall k \geq n \text{ almost surely}$$

since k takes countable values (intersection of countable sets of full measure also has full measure). Now taking limits as $n \to \infty$ we are done. □

Theorem (Conditional Dominated Convergence Theorem)**.** *Suppose $X_n \to X$ almost surely, $|X_n| \leq Y$ almost surely with Y integrable. Then $\mathbb{E}[X_n|\mathcal{G}] \to \mathbb{E}[X|\mathcal{G}]$ almost surely.*

Proof. We apply the Conditional Fatou's Lemma. Indeed $-Y \leq X_n \leq Y$ so $X_n + Y \geq 0$ and $Y - X_n \geq 0$ for all n. By Conditional Fatou's Lemma

$$\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[X+Y|\mathcal{G}] = \mathbb{E}[\liminf_{n\to\infty}(X_n + Y)] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$$

and

$$\mathbb{E}[Y|\mathcal{G}] - \mathbb{E}[X|\mathcal{G}] = \mathbb{E}[\liminf_{n\to\infty}(Y - X_n)|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}] + \liminf_{n\to\infty} \left(-\mathbb{E}[X_n|\mathcal{G}]\right).$$

Hence $\limsup_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \leq \mathbb{E}[X|\mathcal{G}]$ and $\liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \geq \mathbb{E}[X|\mathcal{G}]$ almost surely.

□

Theorem (Conditional Jensen's Inequality)**.** *Let X be integrable, $\varphi : \mathbb{R} \to \mathbb{R}$ a convex function such that $\varphi(X)$ is integrable or $\varphi(X) \geq 0$. Then $\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathbb{E}[X|\mathcal{G}])$ almost surely.*

Proof. We claim that $\varphi(x) = \sup_{i\in\mathbb{N}}(a_i x + b_i)$, $a_i, b_i \in \mathbb{R}$.

Then $\varphi(X) = \sup_{i\in\mathbb{N}}(a_i X + b_i)$. So

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \sup_{n\geq 1}(a_i \mathbb{E}[X|\mathcal{G}] + b_i) \quad \forall i \in \mathbb{N} \text{ almost surely.}$$

□

Note. We need the supremum in the claim to be over a countable set so we can preserve the almost-sue property of an inequality.

Corollary. *For all $p \in [1, \infty)$ we have*

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

Proof. Apply conditional Jensen ($x \mapsto x^p$ is convex). □

Theorem (Tower property)**.** *Let X be integrable and $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ sub σ-algebras. Then*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] \textit{ almost surely.}$$

Proof. $\mathbb{E}[X|\mathcal{H}]$ is certainly $\mathcal{H}$-measurable so it remains to check

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]\mathbb{1}(A)] \quad \forall A \in \mathcal{H}.$$

But since $A \in \mathcal{G}$ whenever $A \in \mathcal{H}$ we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}(A)] = \mathbb{E}[X\mathbb{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]\mathbb{1}(A)].$$

□

Proposition. Let $X \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$ a sub σ-algebra, Y bounded and $\mathcal{G}$-measurable. Then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \text{ almost surely.}$$

Proof. $Y\mathbb{E}[X\mathcal{G}]$ is certainly $\mathcal{G}$-measurable. Also for any $A \in \mathcal{G}$

$$\mathbb{E}[XY\mathbb{1}(A)] = \mathbb{E}[X \underbrace{(Y\mathbb{1}(A))}_{\substack{\text{bounded,}\\ \mathcal{G}\text{-measurable}}}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}](Y\mathbb{1}(A))].$$

□