Part III Advanced Probability

Based on lectures by P. Sousi

Notes taken by Pantelis Tassopoulos

Michaelmas 2023

Contents

1 Conditional Expectation

Lecture 1

1.1 Basic definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Remember the following definitions

Definition 1.1 (Sigma algebra). $\mathcal{F}$ is a sigma algebra if and only if: $(\mathcal{F} \in \mathcal{P}\Omega)$

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
3. $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{F} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

Definition 1.2 (Probability measure). $\mathbb{P}$ is a probability measure if

1. $\mathbb{P} : \mathcal{F} \to [0,1]$ (i.e. a set function)
2. $\mathbb{P}(\Omega) = 1$, and $\mathbb{P}(\varnothing) = 0$
3. $(A_n)_{n\in\mathbb{N}}$ pairwise disjoint $\implies \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Definition 1.3 (Random Variable). $X : \Omega \to \mathbb{R}$ is a <u>random variable</u> if for all B open in $\mathbb{R}$, $X^{-1}(B) \in \mathcal{F}$.

Remark. Observe that the sigma algebra $\mathcal{G} = \{B \subseteq \mathbb{R} : X(B) \in \mathcal{F}\} \supseteq \mathcal{O} \implies \mathcal{G} \supseteq \mathcal{B}(\mathbb{R})$, the former being the collection of open sets in $\mathbb{R}$ and the latter the Borel sigma algebra on $\mathbb{R}$ with the usual topology, namely, $\sigma(\mathcal{O})$ (see below for the notation).

Let $\mathcal{A}$ be a collection of subsets of Ω. We define

$$\begin{aligned}\sigma(\mathcal{A}) &= \text{smallest sigma algebra containing } \mathcal{A}\\ &= \bigcap\{\mathcal{T} : \mathcal{T} \text{ sigma algebra containing } \mathcal{A}\}.\end{aligned}$$

Definition 1.4 (Borel sigma algebra on $\mathbb{R}$). Let $\mathcal{O} = \{\text{open sets}\mathbb{R}\}$. Then, the Borel sigma algebra $\mathcal{B}(\mathbb{R})(\coloneqq \mathcal{B})$ is defined as above, namely,

$$\mathcal{B}(\mathbb{R}) \coloneqq \sigma(\mathcal{O}).$$

Let $(X_i)_{i\in I}$ be a family of random variables, then $\sigma(X_i : i \in I) =$ the smallest sigma algebra that makes them all measurable. We also have the characterisation $\sigma(X_i : i \in I) = \sigma(\{\underbrace{\{\omega \in \Omega : X_i(\omega) \in B\}}_{X_i^{-1}(B)}, i \in I, B \in \mathcal{B}(\mathbb{R})\})$.

1.2 Expectation

Note we use the following for the indicator function on some event A

$$\mathbf{1}(A)(x) = \mathbf{1}(x \in A) \coloneqq \left.\begin{matrix} 1, & x \in A \\ 0, & x \notin A \end{matrix}\right\}, \quad A \in \mathcal{F}.$$

We now begin the construction of the expectation of generic random variables.

<u>Positive simple random variables:</u> $X = \sum_{i=1}^{n} \mathbf{1}(A_i), c_i \geqslant 0, A_i \in \mathcal{F}$..

$$\mathbb{E}[X] \coloneqq \sum_{i=1}^{n} c_i \mathbb{P}(A_i).$$

Non-negative random variables: ($X \geqslant 0$). We proceed by approximation. Namely, let $X_n(\omega) := 2^{-n}\lfloor 2^{-n} \cdot X(\omega) \rfloor \wedge n \uparrow X(\omega), n \to \infty$. Now, by monotone convergence,

$$\mathbb{E}[X] :=\uparrow \lim_{n\to\infty} \mathbb{E}[X_n] = \sup \mathbb{E}[X].$$

General random variables: Have the decomposition $X = X^+ - X^-$, where $X^+ = X \vee 0$, $X^- = -X \wedge 0$. If one of $\mathbb{E}[X^+], \mathbb{E}[X^-] < \infty$ then set

$$\mathbb{E}[X] := \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

Definition 1.5. X is called integrable if $\mathbb{E}[|X|] < \infty$.

Definition 1.6. Let $B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$. Then for all $A \in \mathcal{F}$, set

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

Now for an integer-valued random variable X, we set:

$$\mathbb{E}[X|B] := \frac{\mathbb{E}[X \cdot \mathbf{1}_B]}{\mathbb{P}(B)}$$

1.3 Conditional expectation with respect to countably generated sigma algebras

Lecture 2 We now extend the definition of the conditional expectation for a countably generated sigma algebra. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We call the sigma algebra $\mathcal{G}$ coutnably generated if there exists a colection $(B_n)_{n\in\mathbb{N}}$ of pairwise disjoint events such that $\bigcup_{n\in I} B_n = \Omega$ with (I countable) and $\mathcal{G} = \sigma(B_i : i \in I)$.

Let X be an integrable random variable. We want to define $\mathbb{E}[X|\mathcal{G}]$.

Define $X'(\omega) = \mathbb{E}[X|B_i]$, whenever $w \in B_i$, i.e.

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

We make the convention that $\mathbb{E}[X|B_i] = 0$ if $\mathbb{P}(B_i) = 0$. It is easy to check that X' is $\mathcal{G}-$measurable. We also have that

$$\mathcal{G} = \left\{ \bigcup_{j\in} B_j : J \subseteq I \right\}$$

and X' satisfies for all $G \in \mathcal{G}$:$\mathbb{E}[X \cdot \mathbf{1}_G] = \mathbb{E}[X' \cdot \mathbf{1}_G]$ and

$$\begin{aligned}
\mathbb{E}[|X'|] &\leqslant \mathbb{E}\left[\sum_{i\in I} |\mathbb{E}[X|B_i]\mathbf{1}(B_i)\right] \\
&= \sum_{i\in I} \mathbb{P}(B_i) \cdot |\mathbb{E}[X|B_i]| \\
&\leqslant \sum_{i\in I} \mathbb{P}(B_i) \cdot \underbrace{\mathbb{E}[X \cdot \mathbf{1}(B_i)]}_{\mathbb{P}(B_i)} \\
&= \mathbb{E}[|X|] < \infty.
\end{aligned}$$

1.4 General case

We say $A \in \mathcal{F}$ happens a.s. if $\mathbb{P}(A) = 1$. Recall (from measure theory and basic functional analysis):

Theorem 1.1 (Monotone Convergence Theorem (MCT)). Let $(X_n)_{n\in\mathbb{N}}$ be such that $X_n \geqslant 0, X$ be random variables such that $X_n \uparrow X$ as $n \to \infty$. Then, $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$ as $n \to \infty$.

Theorem 1.2 (Dominanted Convergenec Theorem (DCT)). Let $(X_n)_{n\in\mathbb{N}}$ be random variables such that $X_n \to X$ a.s. as $n \to \infty$ and $|X_n| \leqslant Y$ a.s. for all $n \in \mathbb{N}$, where Y is integrable, then $\mathbb{E}[X_n] \to \mathbb{E}[X]$, as $n \to \infty$.

Let $1 \leqslant p < \infty$ and f a measurable function, then set $\|f\|_p := (\mathbb{E}[\|f\|^p])^{\frac{1}{p}}$. If $p = \infty$, then set $\|f\|_\infty := \inf\{\lambda : |f| \leqslant \lambda \text{ a.s.}\}$. Recall for all p, the Lebesgue spaces, $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \|f\|_p < \infty\}$.

Theorem 1.3. $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space, with inner product $\langle u, v\rangle_2 = \mathbb{E}[u \cdot v]$. Furthermore, for any closed subspace $\mathcal{H}$, if $f \in \mathcal{L}^2$, there exists a unique $g \in \mathcal{H}$ s.t. $\|f - g\|_{\mathcal{L}^2} = \inf_{h\in\mathcal{H}} \|f - h\|_{\mathcal{L}^2}$ and $\langle f - g, h\rangle = 0$, for all $h \in \mathcal{H}$. We say that g is the <u>orthogonal projection</u> of f in $\mathcal{H}$.

We now construct the conditional expectation in the general case, for any integrably random variable with respect to an arbitrary sigma algebras.

Theorem 1.4 (Conditional Expectation). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ a sub-sigma algebra, $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists an integrable random variable Y satisfying:

1. Y is $\mathcal{G}-$measurable
2. for all $G \in \mathcal{G}, \mathbb{E}[X \cdot \mathbf{1}(G)] = \mathbb{E}[Y \cdot \mathbf{1}(G)]$.

Moreover, Y unique in the sense that if Y' also satisfies the above $1), 2)$, then $Y = Y'$ a.s.. We call Y a version of the conditional expectation of X given G. We write $Y = \mathbb{E}[X\mathcal{G}]$ a.s. If $\mathcal{G} = \sigma(Z)$, where Z is a random variable, then we write $\mathbb{E}[Z] = \mathbb{E}[X|\mathcal{G}]$.

Remark. 2) could be replaced by $\mathbb{E}[X \cdot Z] = \mathbb{E}[Y \cdot Z]$ for all Z bounded $\mathcal{G}-$measurable random variables.

We now state and prove the main theorem of this section:

Proof. (Theorem 1.4) <u>Uniqueness:</u> Let Y, Y' satisfy $1), 2)$. Let $A = \{Y > Y'\} \in \mathcal{G}$. Then 2)

$$\begin{aligned}
&\implies \mathbb{E}[Y \cdot \mathbf{1}(A)] = \mathbb{E}[Y' \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)] \\
&\implies \mathbb{E}[(Y - Y') \cdot \mathbf{1}(A)] = 0 \\
&\implies \mathbb{P}(A) = \mathbb{P}(Y > Y') = 0 \\
&\implies Y \leqslant Y' \text{ a.s..}
\end{aligned}$$

We similarly obtain $Y \geqslant Y'$ a.s., hence we deduce that $Y = Y'$ a.s.

<u>Existence:</u> three steps.

1. Assume that $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Observe that $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace of $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Hence, Theorem 1.3, we have the decomposition $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) = \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \oplus \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$. Then, we have the corresponding decomposition $X = Y + Z$, where $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ and $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ respectively. Define $\mathbb{E}[X\mathcal{G}] := Y$, Y is $\mathcal{G}-$measurable and for all $A \in \mathcal{G}$, $\mathbb{E}[X \cdot \mathbf{1}(A)]\mathbb{E}[Y \cdot \mathbf{1}(A)] = \mathbb{E}[Z \cdot \mathbf{1}(A)]$ since $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$.

   <u>Claim:</u> If $X \geqslant 0$, a.s. then $Y \geqslant 0$ a.s. Indeed, let $A = \{Y < 0\} \in \mathcal{G}$. Then observe that $0 \leqslant \mathbb{E}[X \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)] \leqslant 0$. Hence $\mathbb{E}[Y \cdot \mathbf{1}(A)] = 0$ and so $\mathbb{P}(A) = 0$, gibing $Y = 0$ a.s.

2. Assume $X \geqslant 0$.
   Define $X_n = X \wedge n \leqslant n$, meaning X_n is bounded for all $n \in \mathbb{N}$. So $X_n \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Let $Y_n = \mathbb{E}[X_n]$ a.s.. $(X_n)_{n\in\mathbb{N}}$ is an increasing sequence. By the claim abose, so is $(Y_n)_{n\in\mathbb{N}}$ a.s. Define $Y = \limsup_n Y_n$ meaning Y is $\mathcal{G}-$measurable and $Y =\uparrow \lim_{n\to\infty} Y_n$ a.s. Now, we have that for all $A \in \mathcal{G}$, $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y_n \cdot \mathbf{1}(A)]$. Thus, by theorem 1.1 (MCT), $\mathbb{E}[X \cdot \mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)]$.

3. X general in $\mathcal{L}^1$.
   Decompose as before $X = X^+ - X^-$. Define, $\mathbb{E}[X\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$.

□

Lecture 3

Remark. From the second step of the proof of Theorem 1.4 we see that we can define $\mathbb{E}[X|\mathcal{G}]$ for all $X \geqslant 0$, not necessarily integrable. It satisfies all conditions $1), 2)$ except for the integrability one.

Definition 1.7. $\underbrace{\mathcal{G}_1, \mathcal{G}_2, \dots}_{\text{sigma algebras}} \subset \mathcal{F}$. We call them independent if whenever $G_i \in \mathcal{G}_i$ and $i_1 < \dots i_k$ for some $k \in \mathbb{N}$, then $\mathbb{P}(G_{i_1} \cap \dots \cap G_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(G_{i_j})$.

Moreover, let X be a random variable and $\mathcal{G}$ a sigma algebra, then they are said to be int if $\sigma(X)$ is independent of $\mathcal{G}$.

Properties of conditional expectations: Fix $X, y \in \mathcal{L}^1$, $G \in \mathcal{F}$.

1. $\mathbb{E}[\mathbb{E}[X\mathcal{G}]] = \mathbb{E}[X]$ (take $A = \Omega$)
2. If X is $\mathcal{G}-$measurable, then $\mathbb{E}[X\mathcal{G}] = X$ a.s.
3. If X is independent of $\mathcal{G}$, then $\mathbb{E}[X\mathcal{G}] = \mathbb{E}[X]$
4. If $X \geqslant 0$ a.s., then $\mathbb{E}[X\mathcal{G}] \geqslant 0$ a.s.
5. For $\alpha, \beta \in \mathbb{R}$ $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$
6. $\mathbb{E}[X|\mathcal{G}]| \leqslant \mathbb{E}[|X||\mathcal{G}]$ a.s.

Below we provide expensions of useful measure theoretic results for the expectation to their corresponding conditional counetparts. First recall:

Lemma 1.1 (Fatou's Lemma). Let $X_n \geqslant 0$ for all $n \in \mathbb{N}$. Then

$$\mathbb{E}[\liminf_n X_n] \leqslant \liminf_n \mathbb{E}[X_n] \quad \text{a.s}$$

Theorem 1.5 (Jensen's Inequality). If X is integrable and $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$\phi(\mathbb{E}[X]) \leqslant \mathbb{E}[\phi(X)] \quad \text{a.s.}$$

Now the results themselves:

Theorem 1.6 (Conditional Monotone Convergence theorem (MCT)). Let $\mathcal{G} \subset \mathcal{F}$ be sigma algebras, $X_n \geqslant 0$ a.a. and $X_n \uparrow X$, as $n \to \infty$ a.s. Then

$$\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

Proof. Theorem 1.6 Set $Y_n = \mathbb{E}[X_n\mathcal{G}]$ a.s. Observe that Y_n is a.s. increasing. Set $Y = \limsup_n Y_n$. Y_n is $\mathcal{G}-$measurable, hence, so is Y (as a $\limsup$ of $\mathcal{G}-$measurable random variables) is also $\mathcal{G}-$measurable. Also, $Y = \lim_{n\to\infty} Y_n$ a.s.

Need to show: $\mathbb{E}[Y \cdot \mathbf{1}(A)]\mathbb{E}[X \cdot \mathbf{1}(A)]$ for all $A \in \mathcal{G}$. Indeed,

$$\begin{aligned}\mathbb{E}[Y \cdot \mathbf{1}(A)] &= \mathbb{E}[\lim_{n\to\infty} Y_n \cdot \mathbf{1}(A)] \overset{\text{MCT}}{=} \lim_{n\to\infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] \\ &= \lim_{n\to\infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)].\end{aligned}$$

□

Proof. Theorem 1.1 $\liminf_n X_n = \lim_{n\to\infty}\left(\inf_{k\geqslant n} X_k\right)$, the limit of an increasing sequence. By Theorem 1.1, we have

$$\lim_{n\to\infty} \mathbb{E}[\inf_{k\geqslant n} X_n|\mathcal{G}] = \mathbb{E}[\liminf_n X_n|\mathcal{G}]$$

and

$$\mathbb{E}[\inf_{k\geqslant n} X_k|\mathcal{G}] \overset{\text{a.s.}}{\leqslant} \inf_{k\geqslant n} \mathbb{E}[X_k|\mathcal{G}]^1$$

which gives the result

$$\mathbb{E}[\liminf_n X_n] \leqslant \liminf_n \mathbb{E}[X_n] \quad \text{a.s.}$$

□

Theorem 1.7 (Conditional Dominated Convergence Theorem). SUppose $X_n \to X$ a.s. $n \to \infty$ and $|X_n| \leqslant Y$ a.s. for all $n \in \mathbb{N}$ with Y in tegrable. Then $\mathbb{E}[X_n\mathcal{G}] \to \mathbb{E}[X\mathcal{G}]$ a.s. as $n \to \infty$.

Proof. From $-Y \leqslant X_n \leqslant Y$, we have $X_n + Y \geqslant 0$ for all $n \in \mathbb{N}$ and $Y - X_n \geqslant 0$a.s. By Theorem 1.1,

$$\begin{aligned}\mathbb{E}[X + Y\mathcal{G}] &= \mathbb{E}[\liminf_n (X_n + Y)|\mathcal{G}] \\ &\leqslant \liminf_n \mathbb{E}[X_n + Y|\mathcal{G}] = \liminf_n \mathbb{E}[X_n\mathcal{G}] + \mathbb{E}[X]\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{E}[|X - Y||\mathcal{G}] &= \mathbb{E}[Y - \liminf_n X_n|\mathcal{G}] \\ &\leqslant \mathbb{E}[Y] + \liminf_n \mathbb{E}[X_n|\mathcal{G}]\end{aligned}$$

Hence,

$$\limsup_n \mathbb{E}[X_n|\mathcal{G}] \leqslant \mathbb{E}[X|\mathcal{G}]$$

and

$$\liminf_n \mathbb{E}[X_n|\mathcal{G}] \geqslant \mathbb{E}[X|\mathcal{G}]$$

a.s., concluding the proof.

□

Theorem 1.8 (Conditional Jensen). Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function s.t. $\phi(X)$ is integrable or $\phi(X) \geqslant 0$

$$\phi(\mathbb{E}[X|\mathcal{G}]) \leqslant \mathbb{E}[\phi(X)|\mathcal{G}] \quad \text{a.s.}$$

Proof. Claim: (true for any convex function, no proof given) $\phi(x) = \sup_{i\in\mathbb{N}}(a_i x + b_i)$, $a_i b_i \in \mathbb{R}$. Thus,

$$\mathbb{E}[\phi(X)|\mathcal{G}] \geqslant a_i\mathbb{E}[X|\mathcal{G}] + b_i \quad \text{for all } i \in \mathbb{N}.$$

Taking the supremum gives [2]

$$\begin{aligned}\mathbb{E}[\phi(X)|\mathcal{G}] &\geqslant \sup_{i\in\mathbb{N}} (a_i\mathbb{E}[X|\mathcal{G}] + b_i) \\ &= \phi(\mathbb{E}[X|\mathcal{G}]) \quad \text{a.s.}\end{aligned}$$

□

Corollary 1.8.1. For all $1 \leqslant p < \infty$ $\|\mathbb{E}[X|\mathcal{G}]\|_p \leqslant \|X\|_p$.

Proof. Apply conditional Jensen. □

[1] can take the infinum due to countability that preserves a.s.
[2] can take the supremum due to countability which again preserves a.s.

Proposition 1.1 (Tower Property). Let X be integrable and $\mathcal{H} \subseteq \mathcal{G}$ sigma algebras. Then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] \quad \text{a.s.}$$

Proof. (a) $\mathbb{E}[X|\mathcal{H}]$ is $\mathcal{H}-$measurable.

(b) For all $A \in \mathcal{H}$ NTS:

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

Indeed, both terms above are equal to $\mathbb{E}[X \cdot \mathbf{1}(A)]$ since $A \in \mathcal{G} \subseteq \mathcal{H}$.

□

Proposition 1.2. Let $X \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$, Y bounded $\mathcal{G}-$measurable. Then

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

Proof. (a) RHS is clearly $\mathcal{G}-$measurable.

(b) For all $A \in \mathcal{G}$:

$$\begin{aligned} \mathbb{E}[X \cdot Y \cdot \mathbf{1}(A)] &= \mathbb{E}[Y \cdot \mathbb{E}[X\mathcal{G}] \cdot \mathbf{1}(A)] \\ \mathbb{E}[X \cdot (\underbrace{Y \cdot \mathbf{1}(A)}_{\mathcal{G}\text{-meas. and bounded}})] &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot Y \cdot \mathbf{1}(A)] = RHS. \end{aligned}$$

(Also observe that by a monotone class argument, we have for any integrable function $f : \Omega \to \mathbb{R}$, $\mathbb{E}[X \cdot f] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot f]$) □