Probability Theory
17. Introduction to conditional expectation

Peter Pfaffelhuber

June 25, 2024

Some elementary calculations

- For $X \in \mathcal{L}^1$ and $A, G \in \mathcal{A}$ let

$$E[X|G] := \frac{E[X; G]}{P(G)}, \qquad P(A|G) := \frac{P(A \cap G)}{P(G)}$$

the conditional probability and conditional expectation.

- Goal: Define $E[X|\mathcal{G}]$ for $\mathcal{G} \subseteq \mathcal{A}$ σ-algebra.
- Let $\mathcal{H} = \{G_1, G_2, \dots\} \subseteq \mathcal{F}$ be a partition of Ω and $\mathcal{G} = \sigma(\mathcal{H})$,

$$E[X|\mathcal{G}](\omega) := \sum_{i=1}^{\infty} E[X|G_i] \cdot 1_{G_i}(\omega).$$

Further for $J \subseteq \mathbb{N}$ and $A = \bigcup_{j \in J} G_j \in \mathcal{G}$

$$E[E[X|\mathcal{G}]; A] = E\Big[\sum_{i=1}^{\infty} E[X|G_i] 1_{G_i} 1_A\Big] = \sum_{j \in J} E\big[E[X|G_j] 1_{G_j}\big]$$

$$= \sum_{j \in J} E[X; G_j] = E[X; A].$$

Random success probability

- Example: $U \sim U([0,1])$; given U let $X \sim B(n,U)$. Then
$$P(X = k|U) = \binom{n}{k} U^k (1-U)^{n-k}.$$
Note that
$$E[X|\{U < 1/2\}] = 2E[X1_{U<1/2}] = 2\int_0^{1/2} \sum_{k=0}^{n} k \binom{n}{k} u^k (1-u)^{n-k} du$$
$$= 2\int_0^{1/2} nudu = \tfrac{1}{4}n$$
or
$$E[X|U] = nU.$$

Defining property of conditional expectation

- Theorem 11.2: Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. Then there exists an almost surely unique linear operator $\mathsf{E}[.|\mathcal{G}] : \mathcal{L}^1 \to \mathcal{L}^1$ such that $\mathsf{E}[X|\mathcal{G}]$ is for all $X \in \mathcal{L}^1$ a $\mathcal{G}$-measurable random variable with

  1. $\mathsf{E}\big[\mathsf{E}[X|\mathcal{G}]; A\big] = \mathsf{E}[X; A]$ for all $A \in \mathcal{G}$.

  Proof for $X \in \mathcal{L}^2$: Let $M := \{Y \in \mathcal{L}^2 : \mathcal{G}\text{-measurable}\}$ linear. There are as unique $Y \in M, Z \perp M$ with $X = Y + Z$. Set $\mathsf{E}[X|\mathcal{G}] := Y$. Then, $X - \mathsf{E}[X|\mathcal{G}] \perp M$, therefore

$$\mathsf{E}[X - \mathsf{E}[X|\mathcal{G}]; A] = 0, \qquad A \in \mathcal{G}.$$

Defining property of the conditional expectation

- Theorem 11.2:
  1. $\mathrm{E}\big[\mathrm{E}[X|\mathcal{G}]; A\big] = \mathrm{E}[X; A]$ for all $A \in \mathcal{G}$.
  2. $\mathrm{E}[X|\mathcal{G}] \geq 0$ if $X \geq 0$.
  3. $\mathrm{E}\big[|\mathrm{E}[X|\mathcal{G}]|\big] \leq \mathrm{E}[|X|]$.
  4. If $0 \leq X_n \uparrow X$ for $n \to \infty$, then also $\mathrm{E}[X_n|\mathcal{G}] \uparrow \mathrm{E}[X|\mathcal{G}]$ in $\mathcal{L}^1$.

Proof: 3. With $A := \{\mathrm{E}[X|\mathcal{G}] \geq 0\} \in \mathcal{G}$,

$$\mathrm{E}[|\mathrm{E}[X|\mathcal{G}]|] = \mathrm{E}[\mathrm{E}[X|\mathcal{G}]; A] - \mathrm{E}[\mathrm{E}[X|\mathcal{G}]; A^c] = \mathrm{E}[X; A] - \mathrm{E}[X; A^c] \leq \mathrm{E}[|X|].$$

2. With $A = \{\mathrm{E}[X|\mathcal{G}] \leq 0\} \in \mathcal{G}$,

$$0 \geq \mathrm{E}[\mathrm{E}[X|\mathcal{G}]; A] = \mathrm{E}[X; A] \geq 0.$$

4. Due to monotone convergence, $||X_n - X||_1 \xrightarrow{n\to\infty} 0$, also

$$\mathrm{E}[|\mathrm{E}[X_n|\mathcal{G}] - \mathrm{E}[X|\mathcal{G}]|] = \mathrm{E}[|\mathrm{E}[X_n - X|\mathcal{G}]|] \leq \mathrm{E}[|X_n - X|] \xrightarrow{n\to\infty} 0.$$

Defining property of the conditional expectation

- Theorem 11.2:
  1. $\mathrm{E}\big[\mathrm{E}[X|\mathcal{G}];A\big] = \mathrm{E}[X;A]$ for all $A \in \mathcal{G}$.
  5. X is $\mathcal{G}$-measurable $\Rightarrow \mathrm{E}[XY|\mathcal{G}] = X\mathrm{E}[Y|\mathcal{G}]$.
  6. $\mathrm{E}\big[X\mathrm{E}[Y|\mathcal{G}]\big] = \mathrm{E}\big[\mathrm{E}[X|\mathcal{G}]Y\big] = \mathrm{E}\big[\mathrm{E}[X|\mathcal{G}]\mathrm{E}[Y|\mathcal{G}]\big]$.
  7. If $\mathcal{H} \subseteq \mathcal{G}$, then $\mathrm{E}\big[\mathrm{E}[X|\mathcal{G}]|\mathcal{H}\big] = \mathrm{E}[X|\mathcal{H}]$.
  8. If X is independent of $\mathcal{G}$, then $\mathrm{E}[X|\mathcal{G}] = \mathrm{E}[X]$.

Proof: 6. for $X, Y \in \mathcal{L}^2$. Then, $\mathrm{E}[Y|\mathcal{G}] \in M$ and

$$\mathrm{E}[(X - \mathrm{E}[X|\mathcal{G}])\mathrm{E}[Y|\mathcal{G}]] = 0.$$

5. $A \in \mathcal{G}$ is $\mathrm{E}[X|\mathcal{G}]1_A = X1_A$, thus $\mathrm{E}[XY;A] = \mathrm{E}[X\mathrm{E}[Y|\mathcal{G}];A]$

7. For $A \in \mathcal{H} \subseteq \mathcal{G}$ is $\mathrm{E}[\mathrm{E}[X|\mathcal{G}];A] = \mathrm{E}[X;A] = \mathrm{E}[\mathrm{E}[X|\mathcal{H}];A]$

8. $A \in \mathcal{G}$ is $\mathrm{E}[\mathrm{E}[X|\mathcal{G}];A] = \mathrm{E}[X;A] = \mathrm{E}[X]\mathrm{E}[1_A] = \mathrm{E}\big[\mathrm{E}[X];A\big]$

Jensen's inequality

- Proposition 11.4: I open interval, $\mathcal{G} \subseteq \mathcal{A}$ and $X \in \mathcal{L}^1$ with values in I and $\varphi : I \to \mathbb{R}$ convex. Then,

$$\mathsf{E}[\varphi(X)|\mathcal{G}] \geq \varphi(\mathsf{E}[X|\mathcal{G}]).$$

Uniform integrability and conditional expectation

- Lemma 11.5: Let $X \in \mathcal{L}^1$. Then, $(\mathrm{E}[X|\mathcal{G}])_{\mathcal{G}\subseteq\mathcal{A}}$ is uniformly integrable.

  Since $\{X\}$ is uniformly integrable, there is $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ monotonically increasing, convex with $\frac{\varphi(x)}{x} \xrightarrow{x\to\infty} \infty$ and $\mathrm{E}[\varphi(|X|)] < \infty$. Thus

  $$\sup_{\mathcal{F}\subseteq\mathcal{A}} \mathrm{E}[\varphi(|\mathrm{E}[X|\mathcal{F}]|)] \leq \mathrm{E}[\varphi(|X|)] < \infty.$$

  This means that $\{\mathrm{E}[X|\mathcal{F}] : \mathcal{F} \subseteq \mathcal{A}\ \sigma\text{-algebra}\}$ uniformly integrable.

Dominated, monotone convergence

- For $\mathcal{G} \subseteq \mathcal{F}$ and $X_1, X_2, \dots \in \mathcal{L}^1$ with
  1. $X_n \uparrow X \in \mathcal{L}^1$ almost surely or
  2. $|X_n| \le Y \in \mathcal{L}^1$ for all n, and $X_n \xrightarrow{n\to\infty} X$ almost surely.

  Then

$$\mathsf{E}[X_n|\mathcal{G}] \xrightarrow{n\to\infty}_{as, L^1} \mathsf{E}[X|\mathcal{G}].$$

$\mathcal{L}^1$-convergence: $\mathsf{E}\Big[\big|\mathsf{E}[X_n|\mathcal{G}] - \mathsf{E}[X|\mathcal{G}]\big|\Big] \le \mathsf{E}[|X_n - X|] \to 0.$

as, 1.: $\mathsf{E}[X_n|\mathcal{G}] \uparrow \sup_n \mathsf{E}[X_n|\mathcal{G}]$ and for $A \in \mathcal{G}$

$$\mathsf{E}\Big[\sup_n \mathsf{E}[X_n|\mathcal{G}]; A\Big] = \sup_n \mathsf{E}\Big[\mathsf{E}[X_n|\mathcal{G}]; A\Big] = \sup_n \mathsf{E}[X_n; A] = \mathsf{E}[X; A].$$

as, 2.: Use monotone convergence for

$$Y_n := \sup_{k\ge n} X_k \downarrow \limsup_n X_n = X, \qquad Z_n := \inf_{k\ge n} X_k \uparrow \liminf_n X_n = X$$