Probability Theory 2

April 28, 2025

Contents

1 Conditional Expectation

1.1 Definitions

Example 1.1 (Naive Conditional Expectation)**.** Rolling a die: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Let $X : \Omega \to \mathbb{R}$ where $X =$ score on the die. Define event $A = \{\text{roll is odd}\} = \{X \text{ is odd}\} = \{1, 3, 5\}$

Then conditional expectation of X given event A:

$$E[X \mid A] = \frac{\frac{1}{6}(1+3+5)}{\frac{1}{2}} = \frac{E[X 1_A]}{P(A)} = 3$$

Interpretation: Expectation of X given that A has happened.

Definition 1.2 (Conditional Expectation)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. X is a random variable on Ω, with $E|X| < \infty$ and $\mathcal{G} \subset \mathcal{F}$ is a σ-algebra. We say $Y : \Omega \to \mathbb{R}$ is a version of conditional expectation of X w.r.t. $\mathcal{G}$ (denoted $E[X \mid \mathcal{G}]$) if

1. Y is $\mathcal{G}$-measurable.
2. $\forall B \in \mathcal{G}$, $E[Y 1_B] = E[X 1_B]$.

Interpretation: think of $\mathcal{G}$ as partial information ($\mathcal{G} \subset \mathcal{F}$). We want to replace random variable X by a random variable where information in $\mathcal{G}$ is fixed and the rest is averaged.

We shall use the following notation: $E[Y, A] = E[Y 1_A]$.

Lemma 1.3 (Existance, Uniqueness, Integrability)**.**

(i) Let Y be a version of conditional expectation, then Y is integrable

(ii) Y is unique in the sense that if Y' is another version of condition expectation,then $Y = Y'$ a.s.

(iii) If $E|X| < \infty$, then there exists a version of $E[X \mid \mathcal{G}]$.

Proof. (i) see [Dur Lem 4.1.1].

(ii) If Y and Z are two versions of conditional expectation of X w.r.t. $\mathcal{G}$. Let $\epsilon > 0$, consider the set $A = \{Z - Y \geq \epsilon\}$. Clearly $A \in \mathcal{G}$, then

$$0 = E[X, A] - E[X, A] = E[Z - Y, A] \geq \epsilon P(A)$$

Then we have $P(A) = 0$, which implies $Z \leq Y$ a.s. By the same logic, we also have $Y \leq Z$ a.s. All this leads to

$$Y = Z \quad \text{a.s.}$$

We denote any version of conditional expectation by $E[X \mid \mathcal{G}]$.

(iii) This part uses the Radon-Nikodym Theorem, which states if ν, μ are σ-finite measures, $\nu \ll \mu$, then there exist a function $f = d\nu/d\mu$ (the Radon-Nikodym derivative) such that $\int_A f d\mu = \nu(A)$, for all $A \in \mathcal{F}$. Assume $X \geq 0$ (otherwise use $X = X^+ - X^-$). Consider measure ν on $(\Omega, \mathcal{F}, P)$ defined by

$$\nu(A) = E[X, A], \quad A \in \mathcal{G}.$$

Clearly $\nu(A) = 0$ if $P(A) = 0$. So $\nu \ll P$. Therefore there exist $Y \in L^1(\Omega, \mathcal{G}, P)$ s.t.

$$E[X, A] = \nu(A) = \int_A Y dP = E[Y, A]$$

□

Example 1.4 (Roll of a die)**.** Let X be the outcome of a roll of a die and let event $A = \{X \text{ is odd}\}$. Define $\mathcal{G} = \sigma(A) = \left\{\varnothing, \Omega, A, A^C\right\}$ (information of whether A has happened). We wish to understand $E[X \mid \mathcal{G}]$.

Recall: B is an **atom** of $\mathcal{G}$ if B has no proper subsets in $\mathcal{G}$ except for B.

Claim 1: Let $Y : \Omega \to \mathbb{R}$ be $\mathcal{G}$-measurable, then Y is constant on the atoms of $\mathcal{G}$.

- Let A be an atom and $a \in \mathbb{R}$, $Y^{-1}(\{a\}) \cap A \subseteq A$.
- Since Y is $\mathcal{G}$-measurable, $Y^{-1}(\{a\}) \cap A \in \mathcal{G}$ and cannot be strict subset of A, unless it is $\varnothing$.

Therefore we can assume our conditional expectation has the form:

$$E[X \mid \mathcal{G}] = a 1_A + b 1_{A^c}$$

Then $E[X, A] = E[E[X|A], A] = aE[1_A, A] = aP(A)$. Observe we have recovered the naive definition $E[X \mid A] = E[X, A]/P(A) = a$. We can apply the same argument with $B = A^c$, and then $E[X \mid A^c] = b$.

Example 1.5 (An Extension)**.** The previous example generalizes to a very special but important case:

$$\mathcal{G} = \sigma(\{A_i, i \in \mathcal{I}\})$$

where $\mathcal{I}$ is finite or countable infinite, satisfying $A_i \cap A_j = \varnothing$ for $i \neq j$ and $\bigcup A_i = \Omega$.

Claim 1:

$$\mathcal{G} = \left\{ \bigcup_{i \in I} A_i\,, \quad I \subseteq \mathcal{I} \right\}$$

Just need to show RHS is a σ-algebra.

- $\Omega = \bigcup_{i \in \mathcal{I}} A_N \in \mathcal{G}$.
- $(\bigcup_{i \in I} A_i)^c = \bigcup_{i \notin I} A_i \in \mathcal{G}$
- Let $B_\alpha = \bigcup_{i \in I_\alpha} A_i$ for $\alpha \in \mathbb{N}$. Then $\bigcup_\alpha B_\alpha = \bigcup_\alpha \bigcup_{i \in I_\alpha} A_i = \bigcup_{i \in \cup_\alpha I_\alpha} A_i \in \mathcal{G}$.

Claim 2: $E\left[X \mid \mathcal{G}\right]$ is constant on the atoms of $\mathcal{G}$, which are the A_i's. In other words:

$$E\left[X \mid \mathcal{G}\right](\omega) = \sum_{j \in \mathcal{I}} \alpha_j 1_{A_j}(\omega)$$

- Similar to previous example, $E\left[E\left[X \mid \mathcal{G}\right], A_i\right] = \alpha_i P(A_i) = E\left[X, A\right]$
- If $P(A_i) \neq 0 \implies \alpha_i = E\left[X \mid A_i\right] = \frac{E[X, A_i]}{P(A_i)}$. Then

$$E\left[X \mid \mathcal{G}\right](\omega) = \sum_{j : P(A_j) \neq 0} E\left[X \mid A_j\right] 1_{A_j}(\omega) \quad \text{a.s.}$$

How general is example 1.5?

Suppose $\mathcal{G} = \sigma(B_1, \ldots, B_m)$ where m is finite, then $\mathcal{G}$ is generated by a partition whose atoms are

$$\left\{ \bigcap_{i=1}^{m} C_i \ , \ C_i \in \{B_i, B_i^c\} \right\}$$

This means conditional expectation w.r.t. σ-algebras for discrete problems can be fully understood by example 1.5.

2 Discrete Time Martingale Theory

Defining Martingales

Definition 2.1 (Filtration)**.** A filtration of $(\Omega, \mathbb{F})$ is an increasing sequence $(\mathcal{F}_n)_{n\geq 0}$ of σ-algebras. s.t. for all n

$$\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \cdots \subset \mathcal{F}$$

Definition 2.2 (filtered probability space)**.** Filtered probability space is the tuple $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$.

Definition 2.3 (adapted process)**.** Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$, a process is called adapted if

$$X_n \in \mathcal{F}_n \quad \forall n \geq 0$$

Example 2.4 (natural filtration generated by stochastic process)**.** Given $(\Omega, \mathcal{F}, \mathbb{P})$ and stochastic process $(X_n)_{n\geq 0}$.

$$\mathcal{F}_n^X := \sigma(X_0, X_1, \ldots, X_n)$$

is called the natural filtration generated by X.

Another process (Y_n) is adapted $\iff$ $Y_n = f_n(X_0, \cdots, X_n)$

Definition 2.5 (Martingale)**.** Given $(\Omega, \mathcal{F}, \mathbb{P})$. A stochastic process $(M_n)_{n\geq 0}$ is called sub(super) martingale if $\forall n \in \mathbb{N}$

1. M_n is adapted
2. $E|M_n| < \infty$
3. $M_n \leq (\geq) E[M_{n+1} \mid \mathcal{F}_n]$

Martingale if both super and sub martingale ❖

examples

Example 2.6 (Random Walk)**.**

Properties of Martingales

Proposition 2.7. *Let (M_n) be submartingale w.r.t. $(\mathcal{F}_n)$. Then (M_n) is submartingale w.r.t. $(\mathcal{F}_n^M)$*

Proof. We have $M_n \leq E[M_{n+1} \mid \mathcal{F}_n]$, take conditional expectation on both sides

$$M_n = E[M_n \mid \mathcal{F}_n^M] \leq E\left[E(M_{n+1} \mid \mathcal{F}_n) \mid F_n^M\right] = E[M_{n+1} \mid \mathcal{F}_n]$$

□

Proposition 2.8. *Assume $(M_n)_n$ is martingale and φ convex function, and $\varphi(M_n) \in L^1$, then*

$$\varphi(M_n) \text{ is a submartingale.}$$

Proof. $\varphi(M_n) = \varphi(E[M_{n+1}|M_n]) \leq E[\varphi(M_{n+1})|\mathcal{F}_n]$ (conditional jensen) □

Summary:

- M_n martingale and φ convex $\longrightarrow$ submartingale.
- M_n martingale and φ concave $\longrightarrow$ supermartingale.

Naming of sub and super comes from analysis:

- convex $\longrightarrow$ subharmonic
- concave $\longrightarrow$ superharmonic

Lemma 2.9. *if M_n is submartingale and φ non-decreasing and convex, then*

$$\varphi(M_n) \text{ is submartingale.}$$