Chapter 3: Conditional Expectations

Lecturer: Kenneth Ng

Preview

This chapter develops the theory of conditional expectations, which formalizes how probabilities and expectations adjust in light of new information. We begin by reviewing the notions of conditional probability and independence. Then, we introduce conditional expectation in the measure-theoretic sense and connect it to the familiar concept of conditional expectation given an event from elementary probability theory.

Key topics in this chapter:

1. Conditional probabilities and independence;
2. Conditional expectations;
3. Tower property of conditional expectations.

1 Conditional Probabilities and Independence

We briefly review the concept of conditional probabilities, which describes the likelihood of occurrence of an event A given the occurrence of another event B.

Definition 1.1 Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose that $A, B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$. Then, the ***conditional probability*** of A given B is defined as

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

If the occurrence of B does not change the probability of A, then A and B are said to be *independent*.

Definition 1.2 The events A and B are said to be ***independent*** if

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

The notion of independence can be generalized to any finite collection of sets.

Definition 1.3 Consider a collection of sets $\mathcal{A} = \{A_1, A_2, \ldots, A_n\} \subseteq \mathcal{F}$. The collection is said to be

1. ***pairwise independent*** if, for any $i \neq j$,

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j);$$

2. ***mutually independent*** if, for any sub-collection $\{A_{i_1}, \ldots, A_{i_k}\}$ of $\mathcal{A}$, where $k \leq n$,

$$\mathbb{P}\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} \mathbb{P}(A_{i_j}).$$

The following formulas on conditional probabilities are useful:

1. **(Law of Total Probability)** For any $A \subseteq \Omega$, and any exhaustive and mutually exclusive events $\{B_i\}_{i=1}^n$ with $\mathbb{P}(B_i) > 0$ for $i = 1, 2, \ldots, n$, we have

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A|B_n)\mathbb{P}(B_n).$$

2. **(Bayes Formula)** For any $A, B \subseteq \Omega$ with $\mathbb{P}(A), \mathbb{P}(B) > 0$,

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

The notion of independence can be extended to σ-algebras:

Definition 1.4 Let $\mathcal{F}$ and $\mathcal{G}$ be two σ-algebras. We say that $\mathcal{F}$ and $\mathcal{G}$ are independent if

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

for any $A \in \mathcal{F}$ and $B \in \mathcal{G}$.

Remark 1.1.

1. By definition, the trivial σ-algebra $\mathcal{F}_0 = \{\varnothing, \Omega\}$ is independent of any σ-algebra $\mathcal{F}$: for any $A \in \mathcal{F}$, $\mathbb{P}(A \cap \varnothing) = 0 = \mathbb{P}(A)\mathbb{P}(\varnothing)$, and $\mathbb{P}(A \cap \Omega) = \mathbb{P}(A) = \mathbb{P}(A)\mathbb{P}(\Omega)$.

2. If $\mathcal{F}$ and $\mathcal{G}$ are independent, and $\mathcal{H} \subseteq \mathcal{G}$ is a sub-σ-algebra, then $\mathcal{F}$ and $\mathcal{H}$ are also independent. Indeed, for any $B \in \mathcal{H}$, it also holds that $B \in \mathcal{G}$, and thus $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for any $A \in \mathcal{F}$, thanks to the independence of $\mathcal{F}$ and $\mathcal{G}$.

Using the definition of independence of σ-algebras, we can define the independence of two random variables as follows:

Definition 1.5 The random variables X, Y on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are said to be independent if $\sigma(X)$ and $\sigma(Y)$ are independent σ-algebras.

The following result give a characterization of independence of random variables.

Theorem 1.2 Let X and Y be random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. The following are equivalent:

1. X and Y are independent;
2. For all bounded measurable functions f and g,

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)].$$

2 Conditional Expectations

The conditional expectation of a random variable X describes its expected value given certain information. In this chapter, we will study three forms of conditioning: on an event, on a σ-algebra, and on another random variable.

2.1 Conditional Expectations Given an Event

We first recall the definition of the conditional expectation of X given an event A:

Definition 2.1 Let X be a random variable and A be a measurable event with $\mathbb{P}(A) > 0$. Then, the conditional expectation of X given A, $\mathbb{E}[X|A]$, is defined as

$$\mathbb{E}[X \mid A] = \frac{\mathbb{E}[X \mathbb{1}_A]}{\mathbb{P}(A)}.$$

Suppose X, Y are two $\mathbb{R}$-valued random variables. If $A = \{Y = y\}$, where Y is a random variable and $y \in \mathbb{R}$, we define $\mathbb{E}[X|Y = y]$ as follows:

1. **Discrete case:** If X, Y, and the pair (X, Y) are discrete random variables with respective probability mass functions P_X, P_Y, and joint distribution $P_{X,Y}$, then

$$\mathbb{E}[X|Y = y] = \sum_x x\, P_{X|Y}(x \mid y) = \sum_x x\, \frac{P_{X,Y}(x, y)}{P_Y(y)}.$$

2. **Continuous case:** If X, Y, and (X, Y) are continuous random variables with density functions f_X, f_Y, and $f_{X,Y}$, then

$$\mathbb{E}[X|Y = y] = \int_{-\infty}^{\infty} x\, f_{X|Y}(x \mid y)\, dx = \int_{-\infty}^{\infty} x\, \frac{f_{X,Y}(x, y)}{f_Y(y)}\, dx.$$

2.2 Conditional Expectations Given a σ-Algebra

Recall that a σ-algebra $\mathcal{F}$ represents the available information. If $\mathcal{G} \subseteq \mathcal{F}$, then $\mathcal{G}$ encodes a coarser, or less detailed, information set. The conditional expectation of X given $\mathcal{G}$, which emerges as a random variable, formalizes how we update the expectation X based only on the information available in $\mathcal{G}$.

Definition 2.2 Let X be an $\mathcal{F}$-random variable with $X \in L^1$, and $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-algebra. Then, the conditional expectation of X given $\mathcal{G}$, denoted by $\mathbb{E}[X|\mathcal{G}]$, is the $\mathcal{G}$-measurable random variable such that, for any $A \in \mathcal{G}$,

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A].$$

3 Properties of Conditional Expectations

The following theorem presents the fundamental properties of conditional expectations:

Theorem 3.1 (Properties of Conditional Expectation) Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra.

1. (**Measurability**) If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X|\mathcal{G}] = X$ a.s.
2. (**Independence**) If $\sigma(X)$ and $\mathcal{G}$ are independent, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ a.s.
3. (**Tower Property**) If $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$, then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$$

   In particular, if we take $\mathcal{H} = \mathcal{F}_0$, we have $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.
4. (**Linearity**) If $X, Y \in L^1$ and $a, b \in \mathbb{R}$, then

$$\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}].$$

5. (**Taking Out What's Known**) If Y is $\mathcal{G}$-measurable and $XY \in L^1$, then

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

6. (**Projection Inequality**) For all $X, Y \in L^2$ where Y is $\mathcal{G}$-measurable,

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

7. (**Non-negativity**) If $X \geq 0$ a.s., then $\mathbb{E}[X|\mathcal{G}] \geq 0$ a.s.
8. (**Monotonicity**) If $X \leq Y$ a.s., then $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$ a.s.

Proof.

1. If $Y = X$, $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$ for any $A \in \mathcal{G}$. Since $Y = X$ is $\mathcal{G}$-measurable, by definition, $\mathbb{E}[X|\mathbb{G}] = Y = X$.

2. Let $Y = \mathbb{E}[X]$, which is a constant and thus $\mathcal{G}$-measurable. By independence, for any $A \in \mathcal{G}$, $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X]\mathbb{E}[\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X]\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$. By definition, we have $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.

3. We begin by showing the first equality. To this end, it suffices to show, for any $A \in \mathcal{H}$,

   $$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]\mathbb{1}_A],$$

   since $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]$ is $\mathcal{H}$-measurable. Indeed, since $A \in \mathcal{H}$ and $\mathcal{H} \subseteq \mathcal{G}$, we have $A \in \mathcal{G}$. By the definition of $\mathbb{E}[X|\mathcal{G}]$, we have $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A]$ for any $A \in \mathcal{H}$. Now, using the definition of $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]$, we have, for any $A \in \mathcal{H}$, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]\mathbb{1}_A]$. The claim follows by combining these two equalities.

   The second equality $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$ can be shown by noting that $\mathbb{E}[X|\mathcal{H}]$ is $\mathcal{G}$-measurable and using the first property of the theorem.

4. The linearity is clear by noting

   $$\begin{aligned}\mathbb{E}[(aX + bY)\mathbb{1}_A] &= a\mathbb{E}[X\mathbb{1}_A] + b\mathbb{E}[Y\mathbb{1}_A] \\ &= \mathbb{E}[a\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] + \mathbb{E}[b\mathbb{E}[Y|\mathcal{G}]\mathbb{1}_A] \\ &= \mathbb{E}[(a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}])\mathbb{1}_A],\end{aligned}$$

   for any $A \in \mathcal{G}$.

5. Suppose that $Y = \mathbb{1}_B$, where $B \in \mathcal{G}$. For any $A \in \mathcal{G}$, $A \cap B \in \mathcal{G}$. Using the definition of conditional expectations,

   $$\mathbb{E}[XY\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_{A\cap B}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_{A\cap B}] = \mathbb{E}[Y\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A].$$

   Hence, $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$ if $Y = \mathbb{1}_B$. By linearity, the result also holds if Y is a simple random variable.

   For general random variables Y, the result can be established through a standard three-step procedure: first for simple random variables, then for non-negative random variables, and finally for integrable random variables. At each step, we verify that

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

   for all $A \in \mathcal{G}$. In particular, the second step requires the DCT. The details are omitted.

6. By the tower property and Property 5,

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

7. Let $A = \{\mathbb{E}[X|\mathcal{G}] < 0\}$. Note that $A \in \mathcal{G}$ since $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable. Assume the contrary that $\mathbb{P}(A) > 0$. By definition of conditional expectations, we have $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A]$. Since $X \geq 0$ a.s., we have $\mathbb{E}[X\mathbb{1}_A] \geq 0$. However, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] < 0$, since $\mathbb{P}(A) > 0$ and $\mathbb{E}[X|\mathcal{G}] < 0$ on A. This yields a contraction and we conclude that $\mathbb{P}(A) = 0$.

8. This follows by noting that $Z := X - Y \geq 0$ a.s., followed by using Properties 7 and 4.

$\square$

Property 6 of Theorem 3.1 says that the difference $X - \mathbb{E}[X|\mathcal{G}]$ is *orthogonal* to any square-integrable, $\mathcal{G}$-measurable random variable Y. The following theorem strengthens this geometric observation; see also Figure 1.

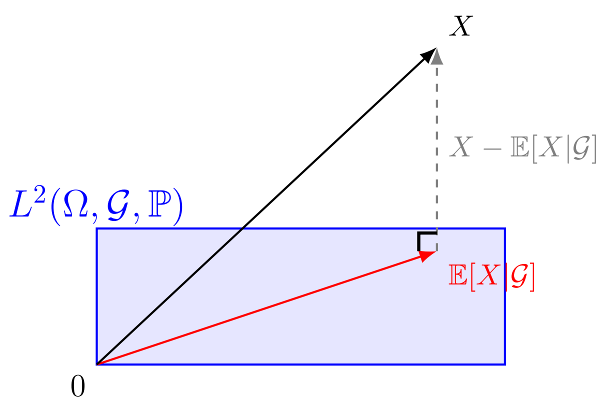

Figure 1: $\mathbb{E}[X|\mathcal{G}]$ is the orthogonal projection of $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ on $L^2(\Omega, \mathcal{G}, \mathbb{P})$

Theorem 3.2 (L^2 **projection of conditional expectations**) The conditional expectation $\mathbb{E}[X|\mathcal{G}]$ is the orthogonal projection of X onto the space of all L^2-integrable, $\mathcal{G}$-measurable random variables. That is, for any $\mathcal{G}$-measurable $Y \in L^2$,

$$\mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2\right] \leq \mathbb{E}\left[(X - Y)^2\right].$$

Proof. For any $\mathcal{G}$-measurable random variable $Y \in L^2$,

$$\begin{aligned}
&\mathbb{E}\left[(X - Y)^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}] + \mathbb{E}[X|\mathcal{G}] - Y)^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2 - 2\left(X - \mathbb{E}[X|\mathcal{G}]\right)\left(Y - \mathbb{E}[X|\mathcal{G}]\right) + \left(Y - \mathbb{E}[X|\mathcal{G}]\right)^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2\right] - 2\mathbb{E}\left[\left(X - \mathbb{E}[X|\mathcal{G}]\right)\left(Y - \mathbb{E}[X|\mathcal{G}]\right)\right] + \mathbb{E}\left[(Y - \mathbb{E}[X|\mathcal{G}])^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2\right] + \mathbb{E}\left[(Y - \mathbb{E}[X|\mathcal{G}])^2\right] \geq \mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G}])^2\right],
\end{aligned}$$

where the second-to-last line follows from Property 6 of Theorem 3.1, by noticing that $Y - \mathbb{E}[X|\mathcal{G}] \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. $\square$

Convergent theorems for expected values introduced in Chapter 2 also apply to conditional expectations. We state the following without proofs:

Theorem 3.3 Let $X \in L^1$ and $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra.

1. (**Jensen's Inequality**) Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function such that $\varphi(X) \in L^1$. Then, $\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$ a.s.
2. (**Monotone Convergence**) Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables such that $X_n \uparrow X$. Then, $\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$ a.s.
3. (**Dominated Convergence**) Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables such that $|X_n| \leq Y$ a.s. for all $n \in \mathbb{N}$, where $Y \in L^1$. Then, $\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$ a.s.

3.1 Relationship of Different Notions

This section relates the conditional expectations given an event, a σ-algebra, and a random variable. The last notion is defined as follows.

Definition 3.1 Let $X \in L^1$. The conditional expectation of X given Y is defined as $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$.

To compute $\mathbb{E}[X|Y]$, one can first determine $\mathbb{E}[X|Y = y]$; see P.3 for the discussions. Let $f(y) := \mathbb{E}[X|Y = y]$. Then, $\mathbb{E}[X|Y] = f(Y)$.

If the σ-algebra $\mathcal{G}$ is generated by a partition $\{A_n\}_{n=1}^{\infty}$ of the sample space Ω, we can determine $\mathbb{E}[X|\mathcal{G}]$ by computing $\mathbb{E}[X|A_n]$ as follows:

Proposition 3.4 Let $\{A_n\}_{n=1}^{\infty}$ be a partition of Ω, i.e., $\cup_n A_n = \Omega$ and $A_n \cap A_m = \varnothing$ for $n \neq m$, such that $\mathbb{P}(A_n) > 0$ for all n. Let $\mathcal{G} := \sigma(\{A_n\}_{n=1}^{\infty})$. Then, for any $X \in L^1$,

$$\mathbb{E}[X|\mathcal{G}] = \sum_{n=1}^{\infty} \mathbb{E}[X|A_n] \mathbb{1}_{A_n}.$$

Proof. Since $A_n \in \mathcal{G}$ for all n, the right-hand side of the formula indeed defines a $\mathcal{G}$-random variable. For any $A \in \mathcal{G}$, we can write $A = \cup_{k=1}^{\infty} A_{n_k}$, each $A_{n_k} \in \{A_n\}_{n=1}^{\infty}$, since the collection is a partition of Ω. Hence,

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}\left[X \sum_{k=1}^{\infty} \mathbb{1}_{A_{n_k}}\right]$$

On the other hand, let $Z := \sum_{n=1}^{\infty} \mathbb{E}[X|A_n]\mathbb{1}_{A_n}$. Then,

$$\mathbb{E}\left[Z\mathbb{1}_A\right] = \mathbb{E}\left[Z \sum_{k=1}^{\infty} \mathbb{1}_{A_{n_k}}\right] = \sum_{k=1}^{\infty} \mathbb{E}[Z\mathbb{1}_{A_{n_k}}]$$

$$
\begin{aligned}
&= \sum_{k=1}^{\infty} \mathbb{E}\left[\left(\sum_{n=1}^{\infty} \mathbb{E}[X|A_n]\mathbb{1}_{A_n}\right)\mathbb{1}_{A_{n_k}}\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}\left[\mathbb{E}[X|A_{n_k}]\mathbb{1}_{A_{n_k}}\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}[X|A_{n_k}]\mathbb{P}(A_{n_k}) \\
&= \sum_{k=1}^{\infty} \mathbb{E}[X\mathbb{1}_{A_{n_k}}] \\
&= \mathbb{E}\left[X\sum_{k=1}^{\infty} \mathbb{1}_{A_{n_k}}\right] = \mathbb{E}[X\mathbb{1}_A],
\end{aligned}
$$

where we have used the DCT and Fubini's theorem when interchanging the infinite summation and integration; the details are omitted herein. The result then follows from the definition of conditional expectations. □

Example 3.1 Consider a two-period binomial model with the sample space $\Omega = \{uu, ud, du, dd\}$, where u and d represents an upward and a downward movement, respectively. The random variable S_2, which represents the value of the equity at the end of the second period, is given by

$$S_2(uu) = 100(1.1)^2,\ S_2(ud) = S_2(du) = 100(1.1)(0.8),\ S_2(dd) = 100(0.8)^2.$$

Suppose that

$$\mathbb{P}(\{uu\}) = p^2,\ \mathbb{P}(\{ud\}) = \mathbb{P}(\{du\}) = p(1-p),\ \mathbb{P}(\{dd\}) = (1-p)^2,$$

where $p \in (0,1)$. Let $X := (S_2 - 80)_+$. Compute $\mathbb{E}[X|\mathcal{G}]$, where $\mathcal{G} := \sigma(U, D)$, $U = \{uu, ud\}$ and $D = \{du, dd\}$.

Solution. Note that $U \cap D = \varnothing$ and $U \cup D = \Omega$. Hence,

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|U]\mathbb{1}_U + \mathbb{E}[X|D]\mathbb{1}_D.$$

On the event U,

$$
\begin{aligned}
\mathbb{E}[X\mathbb{1}_U] &= (100 \times 1.1^2 - 80)_+ p^2 + (100 \times 1.1 \times 0.8 - 80)_+ p(1-p) = p(33p+8), \\
\mathbb{E}[X|U] &= \frac{\mathbb{E}[X\mathbb{1}_U]}{\mathbb{P}(U)} = \frac{p(33p+8)}{p^2 + p(1-p)} = 33p + 8.
\end{aligned}
$$

On the event D,

$$\mathbb{E}[X\mathbb{1}_D] = (100 \times 0.8 \times 1.1 - 80)_+ p(1-p) + (100 \times 0.8^2 - 80)_+ (1-p)^2 = 8p(1-p),$$

$$\mathbb{E}[X|D] = \frac{\mathbb{E}[X\mathbb{1}_D]}{\mathbb{P}(D)} = \frac{8p(1-p)}{p(1-p)+(1-p)^2} = 8p.$$

Therefore,

$$\mathbb{E}[X|\mathcal{G}] = (33p+8)\mathbb{1}_U + 8p\mathbb{1}_D.$$

□

Example 3.2 Let $\Omega = \{a,b,c,d\}$, $\mathcal{F} = 2^\Omega$, and $\mathbb{P}$ be a probability measure that satisfies

$$\mathbb{P}(\{a\}) = \frac{1}{6},\ \mathbb{P}(\{b\}) = \frac{1}{3},\ \mathbb{P}(\{c\}) = \frac{1}{4},\ \mathbb{P}(\{d\}) = \frac{1}{4}.$$

Let X and Y be two random variables given by

$$X(a) = 1,\ X(b) = 1,\ X(c) = -1,\ X(d) = -1,$$
$$Y(a) = 1,\ Y(b) = -1,\ Y(c) = 1,\ Y(d) = -1.$$

Determine $\mathbb{E}[Y|X]$ and verify the partial-averaging property, $\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$.

Solution. Note that $\sigma(X) = \sigma(\{a,b\},\{c,d\})$. Let $A_1 = \{a,b\}$ and $A_2 = \{c,d\}$, we can write

$$\mathbb{E}[Y|X] = \mathbb{E}[Y|\sigma(X)] = \mathbb{E}[Y|A_1]\mathbb{1}_{A_1} + \mathbb{E}[Y|A_2]\mathbb{1}_{A_2}.$$

With $\mathbb{P}(A_1) = \frac{1}{6}+\frac{1}{3} = \frac{1}{2}$, $\mathbb{P}(A_2) = \frac{1}{4}+\frac{1}{4} = \frac{1}{2}$,

$$\mathbb{E}[Y\mathbb{1}_{A_1}] = Y(a)\mathbb{P}(\{a\}) + Y(b)\mathbb{P}(\{b\}) = \frac{1}{6} - \frac{1}{3} = -\frac{1}{6},$$
$$\mathbb{E}[Y|A_1] = \frac{\mathbb{E}[Y\mathbb{1}_{A_1}]}{\mathbb{P}(A_1)} = \frac{-\frac{1}{6}}{\frac{1}{2}} = -\frac{1}{3},$$
$$\mathbb{E}[Y\mathbb{1}_{A_2}] = Y(c)\mathbb{P}(\{c\}) + Y(d)\mathbb{P}(\{d\}) = \frac{1}{4} - \frac{1}{4} = 0,$$
$$\mathbb{E}[Y|A_2] = \frac{\mathbb{E}[Y\mathbb{1}_{A_2}]}{\mathbb{P}(A_2)} = 0.$$

Hence,

$$\mathbb{E}[Y|X] = -\frac{1}{3}\mathbb{1}_{A_1}.$$

Next, we verify the partial-averaging property. Note that

$$\mathbb{E}[Y] = 1\times\frac{1}{6} + (-1)\times\frac{1}{3} + 1\times\frac{1}{4} + (-1)\times\frac{1}{4} = -\frac{1}{6}.$$

On the other hand,

$$\mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}\left[-\frac{1}{3}\mathbb{1}_{A_1}\right] = -\frac{1}{3}\mathbb{P}(A_1) = \frac{1}{6},$$

and so $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$.

□