Note on Conditional Expectation and Conditional Independence

Jiahui Shui

February 10, 2025

1 σ-Algebra

Definition 1. *A collection of subsets $\mathcal{F}$ of Ω is called a σ-algebra (or σ-field) if*

1. $\emptyset, \Omega \in \mathcal{F}$
2. $A \in \mathcal{F}$ *implies* $A^c \in \mathcal{F}$
3. *if* $\{A_i\}_{i=1}^{\infty} \subset \mathcal{F}$, *then* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Let $\mathcal{E}$ be any collection of subsets of Ω. Then there exists a unique smallest σ-algebra that contains $\mathcal{E}$:

$$\sigma(\mathcal{E}) := \bigcap \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma - \text{algebra, and } \mathcal{E} \subset \mathcal{F}\}$$

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We call $X : \Omega \to \mathbb{R}$ a simple random variable if $X(\Omega) \subset \mathbb{R}$ is a finite set. We say that X is $\mathcal{F}$-measurable if $\{X(\omega) = x\} := X^{-1}(\{x\}) \in \mathcal{F}$. More generally,

Definition 2. *We call $X : \Omega \to \mathbb{R}$ a random variable if for all $B \in \mathcal{B}(\mathbb{R})$, $X^{-1}(B) \in \mathcal{F}$. Here $\mathcal{B}(\mathbb{R})$ denote all Borel sets.*

Definition 3 (σ-algebra Generated by Random Variable)**.** *Let X be a random variable, then we call the set:*

$$\sigma(X) = \left\{ X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}) \right\}$$

the σ-algebra generated by X.

Exercise 1. *Prove that $\sigma(X)$ is indeed a σ-algebra where X is a random variable.*

Proof. Note that $X^{-1}(\emptyset) = \emptyset$, hence $\emptyset \in \sigma(X)$. Also, $\forall B \in \mathcal{B}(\mathbb{R})$

$$(X^{-1}(B))^c = \{\omega : X(\omega) \notin B\} = \{\omega : X(\omega) \in B^c\} = X^{-1}(B^c)$$

Finally, for any $\{B_i\}_{i=1}^{\infty} \subset \mathcal{B}(\mathbb{R})$

$$\bigcup_{i=1}^{n} X^{-1}(B_i) = \bigcup_{i=1}^{\infty} \{\omega : X(\omega) \in B_i\} = \left\{ \omega : X(\omega) \in \bigcup_{i=1}^{\infty} B_i \right\} = X^{-1}\left(\bigcup_{i=1}^{\infty} B_i \right)$$

Hence $\sigma(X)$ is a σ-algebra. $\square$

Definition 4 (σ-algebra Generated by Random Variables)**.** *Let $\{X_i\}_{i\in I}$ be random variables, then*

$$\sigma(X_i : i \in I) := \sigma\left(\bigcup_{i\in I} \sigma(X_i)\right)$$

Exercise 2. *Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and f be a $\mathcal{B}(\mathbb{R})$ measurable function. Prove that $f(X)$ is $\sigma(X)$-measurable.*

Proof. Note that

$$\sigma(f(X)) = \{(f \circ X)^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\} = \{X^{-1}(f^{-1}(B)) : B \in \mathcal{B}(\mathbb{R})\}$$

Since f is $\mathcal{B}(\mathbb{R})$-measurable, then $f^{-1}(B) \subset \mathcal{B}(\mathbb{R})$

$$\{X^{-1}(f^{-1}(B)) : B \in \mathcal{B}(\mathbb{R})\} \subset \sigma(X)$$

Thus $\sigma(f(X)) \subset \sigma(X)$ and we finished the proof. □

2 Conditional Expectation

Definition 5. *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$ and let $X \in L^1c$ be a random variable. Then we define $\xi = \mathbb{E}[X|\mathcal{G}]$ as followings:*

(1) $\xi \in L^1$

(2) $\xi \in \mathcal{G}$, *i.e. ξ is $\mathcal{G}$-measurable*

(3) $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\xi\mathbb{1}_A], \forall A \in \mathcal{G}$

In this definition, for any $X \in L^1$, there exists such a conditional expectation. And any two conditional expectations of X are equal $\mathbb{P}$-a.s.

The conditional expectation $\mathbb{E}[X|Y]$ is given by $\mathbb{E}[X|\sigma(Y)]$. Now consider a partition $\{A_k : k \geq 1\}$ of Ω, that is

$$\bigcup_{k\geq 1} A_k = \Omega, \quad \mathbb{P}(A_k) > 0, \forall k \geq 1$$

Also $A_i \cap A_j = \emptyset$ for all $i \neq j$. Let $\mathcal{G} = \sigma(A_k : k \geq 1)$, then

$$\mathbb{E}[X|\mathcal{G}] = \sum_{k=1}^{\infty} \frac{\mathbb{E}[X\mathbb{1}_{A_k}]}{\mathbb{P}(A_k)} \mathbb{1}_{A_k}$$

How to verify this? Let $Y := \sum_{k=1}^{\infty} \frac{\mathbb{E}[X\mathbb{1}_{A_k}]}{\mathbb{P}(A_k)} \mathbb{1}_{A_k}$. Since $A_k \in \mathcal{G}$, then $\forall (\alpha_k)_{k\geq 1} \subset \mathbb{R}$,

$$\sum_{k=1}^{\infty} \alpha_k \mathbb{1}_{A_k} \in \mathcal{G}$$

Hence $Y \in \mathcal{G}$. Now we only have to verify (3) in the definition. $\forall n$

$$\mathbb{E}[Y\mathbb{1}_{A_n}] = \mathbb{E}\left[\sum_{k=1}^{\infty} \frac{\mathbb{E}[X\mathbb{1}_{A_k}]}{\mathbb{P}(A_k)} \mathbb{1}_{A_k}\mathbb{1}_{A_n}\right] = \mathbb{E}[\mathbb{E}[X\mathbb{1}_{A_n}]/\mathbb{P}(A_n)]\mathbb{E}[\mathbb{1}_{A_n}] = \mathbb{E}[X\mathbb{1}_{A_n}]$$

Hence Y is indeed $\mathbb{E}[X|\mathcal{G}]$.

Proposition 1. *Let $X, Y \in L^1$ and let $\mathcal{G}, \mathcal{H}$ be sub-σ-algebra of $\mathcal{F}$, then*

1. *If $X \in \mathcal{G}$, then $X = \mathbb{E}[X|\mathcal{G}]$ a.s.. In particular, $c = \mathbb{E}[c|\mathcal{G}]$ for any constant $c \in \mathbb{R}$.*
2. *(Pull out what's known) If $Y \in \mathcal{G}$ and $XY \in L^1$, then*
$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \tag{1}$$
3. *(Tower property) If $\mathcal{H} \subset \mathcal{G}$, then*
$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{H}] \tag{2}$$
4. *If $\mathcal{H}$ is independent of $\sigma(\mathcal{G}, \sigma(X))$, then*
$$\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X|\mathcal{G}] \tag{3}$$
In particular, if X is independent of $\mathcal{H}$, then $\mathbb{E}[X|\mathcal{H}] = \mathbb{E}[X]$

Proof.
1. By definition this is true.
2. To prove this statement, we have to show that $\forall A \in \mathcal{G}$,
$$\mathbb{E}[XY\mathbb{1}_A] = \mathbb{E}[Y\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A]$$
Consider a more general statement:
$$\mathbb{E}[ZX] = \mathbb{E}[Z\mathbb{E}[X|\mathcal{G}]], \quad \forall Z \in \mathcal{G}, ZX \in L^1$$
Assume that Z and X are both non-negative. Then we can find simple r.v.s $\{Z_n\}$ such that $Z_n \uparrow Z$. By monotone convergence theorem, we must have
$$\mathbb{E}[ZX] = \lim_{n\to\infty} \mathbb{E}[Z_n X] = \lim_{n\to\infty} \mathbb{E}[Z_n \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[Z\mathbb{E}[X|\mathcal{G}]] \tag{4}$$
Note that $Z_n = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{A_k}$, then by definition of conditional expectation and linearity, we will get the second equality. By dominated convergence theorem, we can generalize the statement to $X \in L^1$ instead of only non-negative random variables.
3. Let $Y := \mathbb{E}[X|\mathcal{H}]$. Then $Y \in \mathcal{H} \subset \mathcal{G}$. Then by 1, we know that $\mathbb{E}[Y|\mathcal{G}] = Y$.
On the other hand, let $Z := \mathbb{E}[X|\mathcal{G}]$. Then we only have to show that $\mathbb{E}[Z|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = Y$. $\forall A \in \mathcal{H} \subset \mathcal{G}$,
$$\mathbb{E}[Z\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$$
Then $\mathbb{E}[Z|\mathcal{H}] = Y$ and hence we complete the proof.
4. We will see this in the later section (conditional independence).

□

3 Radon-Nikodym Derivative

4 Conditional Independence