Note: Conditional Expectation and Probability

Atsushi Nitanda

September 5, 2023

1 Conditional expectation regarding a sub-algebra

Let $X : \Omega \to \mathbb{R}$ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra.

Definition 1. *We refer to a random variable Y as a conditional expectation of X under $\mathcal{G}$ when the following are satisfied:*

- *Y is $\mathcal{G}$-measurable.*
- *Y is integrable and satisfies for any $A \in \mathcal{G}$,*

$$\mathbb{E}[X; A] = \int_A X(\omega)P(\mathrm{d}\omega) = \int_A Y(\omega)(\mathrm{d}\omega) = \mathbb{E}[Y; A].$$

We denote by $\mathbb{E}[X|\mathcal{G}]$ a conditional expectation defined above. The existence and almost sure uniqueness can be confirmed by Radon–Nikodým theorem. Let us define the finite measure on $(\Omega, \mathcal{G})$ as follows:

$$\mu(A) = \int_A X(\omega)P(\mathrm{d}\omega) \qquad (A \in \mathcal{G}).$$

Then, μ is absolutely continuous with respect to $P|_{\mathcal{G}}$. That is, if $P|_{\mathcal{G}}(A) = 0$ for $A \in \mathcal{G}$, then $\mu(A) = 0$. By applying Radon–Nikodým theorem, there exists a $\mathcal{G}$-measuarable and P-integrable function Y that satisfies $\int_A X(\omega)P(\mathrm{d}\omega) = \int_A Y(\omega)(\mathrm{d}\omega)$. Moreover, Y which satisfies these conditions is unique P-almost surely.

Example 1. *Let $\{\Omega_j\}_{j\in\mathbb{N}}$ be a partition of Ω. That is, $\Omega = \cup_j \Omega_j$, $\Omega_j \cap \Omega_k = \emptyset$ $(j \neq k)$, $\Omega_j \in \mathcal{F}$, and $\mathbb{P}[\Omega_j] > 0$ $(j \in \mathbb{N})$. Let $\mathcal{G} = \sigma(\{\Omega_j\}_{j\in\mathbb{N}})$. Then, for an integrable random variable X,*

$$\mathbb{E}[X|\mathcal{G}] = \sum_{j\in\mathbb{N}} \frac{\mathbb{E}[X 1_{\Omega_j}]}{\mathbb{P}[\Omega_j]} 1_{\Omega_j} \qquad \textit{a.s.}$$

Especially, for $A \in \mathcal{F}$ and $X = 1_A$,

$$\mathbb{E}[1_A|\mathcal{G}] = \sum_{j\in\mathbb{N}} \frac{\mathbb{P}[A \cap \Omega_j]}{\mathbb{P}[\Omega_j]} \qquad \textit{on } \Omega_j \textit{ a.s.}$$

2 Conditional expectation regarding a measurable

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$ be an integrable $\mathcal{F}$-measurable function, $(\mathcal{T}, \mathcal{B})$ be a measurable space, and $T : \Omega \to \mathcal{T}$ be a $\mathcal{F}/\mathcal{B}$-measurable map. Let P^T denote the push-forward measure $T_\sharp P$.

Definition 2. *We refer to a function $g : \mathcal{T} \to \mathbb{R}$ as a conditional expectation of X under $T = t$ when the following are satisfied:*

- *g is $\mathcal{B}$-measurable.*
- *g is P^T-integrable and satisfies for any $B \in \mathcal{B}$,*

$$\int_{T^{-1}(B)} X(\omega)P(d\omega) = \int_B g(t)P^T(dt).$$

We denote by $\mathbb{E}[X|T = t]$ a conditional expectation g. The existence and uniqueness up to 0-measurable set regarding P^T can be proven in the same way as in the previous section. That is, we define the finite measure on $(\mathcal{T}, \mathcal{B})$ by

$$\mu(B) = \int_{T^{-1}(B)} X(\omega)P(d\omega) \quad (B \in \mathcal{B}).$$

Then, μ is absolutely continuous in P^T. By Radon–Nikodým derivative $d\mu/dP^T(t)$ satisfies the condition.

Example 2. *Let $Y : \Omega \to \mathbb{R}^{d_1}$ and $T : \Omega \to \mathbb{R}^{d_2}$ be a measurable maps defined on the probability space $(\Omega, \mathcal{F}, P)$. Suppose the joint distribution $P^{(Y,T)}$ of (Y, T) has a probability density function $p^{(Y,T)}(y, t)$. We denote by $p^T(t)$ the density of the marginal distribution in T. We set*

$$p(y|t) = \begin{cases} \frac{p^{(Y,T)}(y,t)}{p^T(t)} & (p^T(t) \neq 0), \\ 0 & (p^T(t) = 0). \end{cases}$$

Suppose for a measurable function $h : \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \to \mathbb{R}$, $h(Y, T) : \Omega \to \mathbb{R}$ is integrable. Then, it follows that for $B \in \mathcal{B}(\mathbb{R}^{d_2})$,

$$\begin{aligned} \int_{T^{-1}(B)} h(Y, T)(\omega)P(d\omega) &= \int_{\mathbb{R}^{d_1} \times B} h(y, t)P^{(Y,T)}(dy, dt) \\ &= \int_{\mathbb{R}^{d_1} \times B} h(y, t)p^{(Y,T)}(y, t)dydt \\ &= \int_B \left\{ \int_{\mathbb{R}^{d_1}} h(y, t)p(y|t)dx \right\} p^T(t)dt. \end{aligned}$$

This means

$$\mathbb{E}[h(Y, T)|T = t] = \int_{\mathbb{R}^{d_1}} h(y, t)p(x|t)dx \qquad P^T\text{-a.s.}$$

The next proposition shows that the conditional expectations under a map T and $\sigma(T)$ are equivalent. Here, $\sigma(T)$ is the minimum σ-algebra so that T is measurable, which is $\sigma(T) = \{T^{-1}(B)|B \in \mathcal{B}\}$.

Proposition 1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $(\mathcal{T}, \mathcal{B})$ be a measuralbe space, $T : \Omega \to \mathcal{T}$ be a random variable, and $X : \Omega \to \mathbb{R}$ be a $\mathcal{F}$-measuralbe integrable random variable. Then,*

$$\mathbb{E}[X|\sigma(T)](\omega) = \mathbb{E}[X|T](\omega) \qquad P\text{-a.s. } \omega,$$

where we define $\mathbb{E}[X|T](\omega) = \mathbb{E}[X|T = t]|_{t=T(\omega)}$.

Proof. For any $A \in \sigma(T)$, there exists $B \in \mathcal{B}$ such that $A = T^{-1}(B)$. Then,

$$\begin{aligned}
\int_A \mathbb{E}[X|\sigma(T)](\omega)P(d\omega) &= \int_A X(\omega)P(d\omega) \\
&= \int_{T^{-1}(B)} X(\omega)P(d\omega) \\
&= \int_B \mathbb{E}[X|T = t]|P^T(dt) \\
&= \int_{T^{-1}(B)} \mathbb{E}[X|T = t]|_{t=T(\omega)}P(d\omega) \\
&= \int_A \mathbb{E}[X|T = t]|_{t=T(\omega)}P(d\omega).
\end{aligned}$$

Therefore, we can conclude $\mathbb{E}[X|\sigma(T)] = \mathbb{E}[X|T = t]$ P-a.s. because of the uniqueness of the conditional expectation (i.e., Radon–Nikodým derivative). □

Hence, $\mathbb{E}[X|\sigma(T)]$ is basically a function induced by the composition of functions $T : \Omega \to \mathcal{T}$ and $\mathbb{E}[X|T = \cdot] : \mathcal{T} \to \mathbb{R}$.

3 Conditional Probability

Definition 3. *Let $(\Omega, \mathcal{F}, P)$ be a probability space.*

1. *Let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. The conditional probability of an event $A \in \mathcal{F}$ under $\mathcal{G}$ is defined by $\mathbb{P}[A|\mathcal{G}] = \mathbb{E}[1_A|\mathcal{G}]$.*
2. *Let $(\mathcal{T}, \mathcal{B})$ be a measurable space and $T : \Omega \to \mathcal{T}$ be a $\mathcal{F}/\mathcal{B}$-random variable. The conditional probability of an event $A \in \mathcal{F}$ under the condition $T = t$ is defined by $\mathbb{P}[A|T = t] = \mathbb{E}[1_A|T = t]$.*

For each case, we see the following relationship. (i) For any $B \in \mathcal{G}$, it follows that

$$\mathbb{P}[A \cap B] = \mathbb{E}[1_A; B] = \mathbb{E}[\mathbb{E}[1_A|\mathcal{G}]; B] = \mathbb{E}[\mathbb{P}[A|\mathcal{G}]; B].$$

An example under $\sigma(\{\Omega_j\}_{j\in\mathbb{N}})$ where $\{\Omega_j\}_{j\in\mathbb{N}}$ is a partition of Ω is given in Example 1. (ii) For any $B \in \mathcal{B}$, it follows that

$$\mathbb{P}[A \cap \{T \in B\}] = \mathbb{P}[A \cap T^{-1}(B)] = \int_{T^{-1}(B)} 1_A(\omega)P(d\omega) = \int_B \mathbb{P}[A|T = t]P^T(dt).$$

For $\mathcal{G} = \sigma(T)$, it follows that by Proposition 1 for any $A \in \mathcal{F}$,

$$\mathbb{P}[A|\sigma(T)](\omega) = \mathbb{P}[A|T = t]|_{t=T(\omega)} \qquad P\text{-a.s. } \omega,$$

Example 3. *Let $Y : \Omega \to \mathbb{R}^{d_1}$ and $T : \Omega \to \mathbb{R}^{d_2}$ be measurable maps defined on the probability space $(\Omega, \mathcal{F}, P)$. Suppose $P^{(Y,T)}$ has a probability density function $p^{(Y,T)}(y,t)$. Then for any $A \in \mathcal{B}(\mathbb{R}^{d_1})$ and $B \in \mathcal{B}(\mathbb{R}^{d_2})$, it follows that*

$$\begin{aligned}
\mathbb{P}[Y \in A, T \in B] &= \mathbb{P}[Y^{-1}(A) \cap T^{-1}(B)] \\
&= \int_{T^{-1}(B)} 1_{Y(\omega)\in A} P(d\omega) \\
&= \int_{\mathbb{R}^{d_1} \times B} 1_A(y) P^{(Y,T)}(dy, dt) \\
&= \int_{\mathbb{R}^{d_1} \times B} 1_A(y) p^{(Y,T)}(y,t) dy dt \\
&= \int_{A \times B} p^{(Y,T)}(y,t) dy dt \\
&= \int_B \left\{ \int_A p(y|t) dy \right\} p^T(t) dt.
\end{aligned}$$

Therefore, we see $\mathbb{P}[Y^{-1}(A)|T = t] = \mathbb{P}[Y \in A|T = t] = \int_A p(y|t)dy$. In other words, $p(y|t)$ is a density function of $\mathbb{P}[Y \in \cdot|T = t]$.

Proposition 2. *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ an integrable function. Let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. Then, $\mathbb{E}[X|\mathcal{G}] = \int X(\omega)\mathbb{P}[d\omega|\mathcal{G}]$ a.s.*

Proof. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of the step-function $X_n = \sum_{i=1}^{m_n} a_{n,i} 1_{A_{n,i}}$ $(A_{n,i} \in \mathcal{F})$ such that for any $\omega \in \Omega$, $X_n(\omega) \to X(\omega)$. Then, for any $A \in \mathcal{G}$,

$$\begin{aligned}
\mathbb{E}[X; A] &= \lim_{n\to\infty} \mathbb{E}[X_n; A] \\
&= \lim_{n\to\infty} \sum_{i=1}^{m_n} a_{n,i} \mathbb{E}[1_{A_{n,i}}; A] \\
&= \lim_{n\to\infty} \sum_{i=1}^{m_n} a_{n,i} \mathbb{E}[\mathbb{P}[A_{n,i}|\mathcal{G}]; A] \\
&= \mathbb{E}\left[\lim_{n\to\infty} \sum_{i=1}^{m_n} a_{n,i} \mathbb{P}[A_{n,i}|\mathcal{G}]; A \right] \\
&= \mathbb{E}\left[\int X(\omega)\mathbb{P}[d\omega|\mathcal{G}]; A \right].
\end{aligned}$$

Therefore, $\mathbb{E}[X|\mathcal{G}] = \int X(\omega)\mathbb{P}[d\omega|\mathcal{G}]$ a.s. □

Applying this proposition for the composition $X = f(Y)$ of a measurable map $Y : \Omega \to \mathbb{R}^{d_1}$ and measurable function $f : \mathbb{R}^{d_1} \to \mathbb{R}$, we have

$$\begin{aligned}
\mathbb{E}[f(Y)|\mathcal{G}] &= \int f(Y(\omega))\mathbb{P}[d\omega|\mathcal{G}] \\
&= \int f(y)\mathbb{P}[Y \in dy|\mathcal{G}] \quad \text{a.s.},
\end{aligned}$$

where $\int \cdot \mathbb{P}[Y \in dy|\mathcal{G}]$ represents an integral by the push-forward measure $Y_\sharp\mathbb{P}[\cdot|\mathcal{G}]$.

In the same way, we can show the following. For a measurable map $T : \Omega \to \mathbb{R}^{d_2}$, we have

$$\mathbb{E}[f(Y)|T = t] = \int f(y)\mathbb{P}[Y \in dy|T = t].$$

Especially, when $P^{(Y,T)}$ has the density function $p^{(Y,T)}(y, t)$ as in Example 3,

$$\mathbb{E}[f(Y)|T = t] = \int f(y)p(y|t)dy.$$