CONDITIONAL EXPECTATIONS

1. Introduction

These notes introduce notion of conditional expectation of a random variable and discuss its properties. Aside basic measure-theoretic and probabilistic tools we use here Radon-Nikodym theorem [Monygham, 2018, Theorem 5.1].

2. Existence of conditional expectations

Fix a probability space $(\Omega, \mathcal{F}, P)$.

Theorem 2.1. *Let $X : \Omega \to \mathbb{C}$ be an integrable random variable and $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$. Then there exists $\mathcal{G}$-measurable and integrable function $f : \Omega \to \mathbb{C}$ such that*

$$\int_G X dP = \int_G f dP$$

for every G in $\mathcal{G}$. Moreover, the set of all $\mathcal{G}$-measurable functions having the property described by the system of equations above is

$$\left\{ g : \Omega \to \mathbb{C} \,\middle|\, g \text{ is } \mathcal{G}\text{-measurable and } f(\omega) = g(\omega) \text{ almost surely} \right\}$$

Proof. We define a complex measure $\nu : \mathcal{G} \to \mathbb{C}$ by formula

$$\nu(G) = \int_G X\, dP$$

for $G \in \mathcal{G}$. Since $\nu \ll P_{|\mathcal{G}}$ and by Radon-Nikodym theorem, we derive that there exists a $\mathcal{G}$-measurable function $f : \Omega \to \mathbb{C}$ such that

$$\nu(G) = \int_G f\, dP$$

The last statement is clear and is left for the reader as an exercise. □

Definition 2.2. Let $X : \Omega \to \mathbb{C}$ be an integrable random variable and $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$. Suppose that $f : \Omega \to \mathbb{C}$ is a $\mathcal{G}$-measurable and integrable function $f : \Omega \to \mathbb{C}$ such that

$$\int_G X\, dP = \int_G f\, dP$$

for every G in $\mathcal{G}$. Then f is called *a version of the conditional expectation of X with respect to $\mathcal{G}$*.

No we define important special case.

Definition 2.3. Let $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$. Let $f : \Omega \to \mathbb{C}$ be a $\mathcal{G}$-measurable, integrable function such that

$$P(A \cap G) = \int_G f\, dP$$

for every $G \in \mathcal{G}$. Then f is called *a version of conditional probability of A with respect to $\mathcal{G}$*.

Now that we discuss basic existence and uniqueness results concerning conditional expectation let us introduce some notation. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $X : \Omega \to \mathbb{C}$ be an integrable random variable and $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$. We denote any version of the conditional expectation of X with respect to $\mathcal{G}$ by a symbol

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

and for every set $A \in \mathcal{F}$ we denote by

$$P[A \,|\, \mathcal{G}]$$

any version of conditional probability of A with respect to $\mathcal{G}$. We also often omit the word version and speak about conditional expectation and conditional probabilities. Nevertheless one should always keep in mind that these are $\mathcal{G}$-measurable and integrable functions defined up to sets in $\mathcal{G}$ of probability zero.

3. Properties of conditional expectation

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G}$ be a σ-sublagebra of $\mathcal{F}$.

Theorem 3.1. *Let Y, X, $\{X_n\}_{n\in\mathbb{N}}$ be integrable random variables $\Omega \to \mathbb{C}$. Then the following results hold.*

(1) *If X, Y have real values and $X \leq Y$ almost surely, then $\mathbb{E}[X \,|\, \mathcal{G}] \leq \mathbb{E}[Y \,|\, \mathcal{G}]$ almost surely.*

(2) $\mathbb{E}[a \cdot X + b \cdot Y \,|\, \mathcal{G}] = a \cdot \mathbb{E}[X \,|\, \mathcal{G}] + b \cdot \mathbb{E}[Y \,|\, \mathcal{G}]$ *almost surely for a, $b \in \mathbb{C}$.*

(3) $\big|\mathbb{E}[X \,|\, \mathcal{G}]\big| \leq \mathbb{E}[|X| \,|\, \mathcal{G}]$ *almost surely.*

(4) *If $\{X_n\}_{n\in\mathbb{N}}$ converges almost surely to X and*

$$|X_n| \leq Y,\ |X| \leq Y$$

almost surely for every $n \in \mathbb{N}$, then $\big\{\mathbb{E}[X_n \,|\, \mathcal{G}]\big\}_{n\in\mathbb{N}}$ converges almost surely to $\mathbb{E}[X \,|\, \mathcal{G}]$.

Proof. For the proof of **(1)**. We have

$$\int_G \mathbb{E}[X \,|\, \mathcal{G}]\, dP = \int_G X\, dP \leq \int_G Y\, dP = \int_G \mathbb{E}[Y \,|\, \mathcal{G}]\, dP$$

Since conditional expectations with respect to $\mathcal{G}$ are $\mathcal{G}$-measurable, we deduce that $\mathbb{E}[X \,|\, \mathcal{G}] \leq \mathbb{E}[Y \,|\, \mathcal{G}]$.

Next we prove **(2)**. Pick $a, b \in \mathbb{C}$. We have

$$\int_G \mathbb{E}[a \cdot X + b \cdot Y \,|\, \mathcal{G}]\, dP = \int_G (a \cdot X + b \cdot Y)\, dP = a \cdot \int_G X\, dP + b \cdot \int_G Y\, dP =$$
$$= a \cdot \int_G \mathbb{E}[X \,|\, \mathcal{G}]\, dP + b \cdot \int_G \mathbb{E}[Y \,|\, \mathcal{G}]\, dP = \int_G \big(a \cdot \mathbb{E}[X \,|\, \mathcal{G}] + b \cdot \mathbb{E}[Y \,|\, \mathcal{G}]\big)\, dP$$

for every $G \in \mathcal{G}$. Since conditional expectations with respect to $\mathcal{G}$ is $\mathcal{G}$-measurable, we derive that $\mathbb{E}[a \cdot X + b \cdot Y \,|\, \mathcal{G}] = a \cdot \mathbb{E}[X \,|\, \mathcal{G}] + b \cdot \mathbb{E}[Y \,|\, \mathcal{G}]$.

For **(3)** assume pick $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and

$$\alpha \cdot \mathbb{E}[X \,|\, \mathcal{G}] = \big|\mathbb{E}[X \,|\, \mathcal{G}]\big|$$

Then

$$\big|\mathbb{E}[X \,|\, \mathcal{G}]\big| = \alpha \cdot \mathbb{E}[X \,|\, \mathcal{G}] = \mathbb{E}[\alpha \cdot X \,|\, \mathcal{G}] = \mathbb{E}[\mathrm{re}\,(\alpha \cdot X)\ |\, \mathcal{G}] \leq \mathbb{E}[|\alpha \cdot X| \,|\, \mathcal{G}] = \mathbb{E}[|X| \,|\, \mathcal{G}]$$

almost surely. Thus **(3)** holds.

Finally we prove that **(4)**. Set $Z_n = \sup_{k \leq n} |X_k - X|$. Then Z_n is nonnegative measurable function and $\lim_{n\to+\infty} Z_n = 0$. Moreover, $\big\{Z_n\big\}_{n\in\mathbb{N}}$ is pointwise decreasing and dominated by $2 \cdot |Y|$. Thus by dominated convergence theorem

$$\lim_{n\to+\infty} \int_\Omega Z_n\, dP = 0$$

Next $\big\{\mathbb{E}[Z_n \,|\, \mathcal{G}]\big\}_{n\in\mathbb{N}}$ are measurable, almost surely pointwise decreasing and nonnegative functions. Moreover, we derive that

$$\lim_{n\to+\infty} \int_\Omega \mathbb{E}[Z_n \,|\, \mathcal{G}]\, dP = \lim_{n\to+\infty} \int_\Omega Z_n\, dP = 0$$

and hence

$$\int_\Omega \Big(\lim_{n\to+\infty} \mathbb{E}[Z_n \,|\, \mathcal{G}]\Big)\, dP = 0$$

This implies that $\lim_{n\to+\infty} \mathbb{E}[Z_n \,|\, \mathcal{G}] = 0$ almost surely. By **(1)** and **(3)** we have

$$\sup_{k\geq n} \left|\mathbb{E}[X_k \,|\, \mathcal{G}] - \mathbb{E}[X \,|\, \mathcal{G}]\right| = \sup_{k\geq n} \mathbb{E}\big[|X_k - X| \,\big|\, \mathcal{G}\big] \leq \mathbb{E}[Z_n \,|\, \mathcal{G}]$$

Therefore

$$\lim_{n\to+\infty} \sup_{k\geq n} \left|\mathbb{E}[X_k \,|\, \mathcal{G}] - \mathbb{E}[X \,|\, \mathcal{G}]\right| = 0$$

and hence $\lim_{n\to+\infty} \mathbb{E}[X_n \,|\, \mathcal{G}] = \mathbb{E}[X \,|\, \mathcal{G}]$. □

Theorem 3.2. *Let $X,\, Y : \Omega \to \mathbb{C}$ be random variables such that $X,\, Y \cdot X$ are integrable and Y is $\mathcal{G}$-measurable. Then*

$$\mathbb{E}[Y \cdot X \,|\, \mathcal{G}] = Y \cdot \mathbb{E}[X \,|\, \mathcal{G}]$$

Proof. First note that the result is clear for $Y = \chi_G$ where $G \in \mathcal{G}$ and also for $\mathbb{R}_{>0}$-linear combination of such functions. Next suppose that $Y : \Omega \to \mathbb{C}$ is integrable $\mathcal{G}$-measurable function with nonnegative real values. Then there exists an nondecreasing sequence $\{Y_n\}_{n\in\mathbb{N}}$ of positive combinations of indicator functions of sets in $\mathcal{G}$ that converges to Y. Note that $|Y_n \cdot X| \leq |Y_n| \cdot |X|$ and $|Y_n \cdot \mathbb{E}[X \,|\, \mathcal{G}]| \leq Y \cdot |\mathbb{E}[X \,|\, \mathcal{G}]|$ for $n \in \mathbb{N}$. Then by dominated convergence theorem

$$\int_G \mathbb{E}[Y \cdot X \,|\, \mathcal{G}]\, dP = \int_G Y \cdot X\, dP = \lim_{n\to+\infty} \int_G Y_n \cdot X\, dP = \lim_{n\to+\infty} \int_G Y_n \cdot \mathbb{E}[X \,|\, \mathcal{G}]\, dP = \int_G Y \cdot \mathbb{E}[X \,|\, \mathcal{G}]\, dP$$

for every $G \in \mathcal{G}$. This implies that $\mathbb{E}[Y \cdot X \,|\, \mathcal{G}] = Y \cdot \mathbb{E}[X \,|\, \mathcal{G}]$. Suppose now that $Y : \Omega \to \mathbb{C}$ is a $\mathcal{G}$-measurable and integrable random variable taking real values. We write $Y_+ = \max\{0, Y\}$ and $Y_- = \min\{0, Y\}$. Then

$$\mathbb{E}[Y \cdot X \,|\, \mathcal{G}] = \mathbb{E}[Y_+ \cdot X \,|\, \mathcal{G}] + \mathbb{E}[Y_- \cdot X \,|\, \mathcal{G}] = Y_+ \cdot \mathbb{E}[X \,|\, \mathcal{G}] + Y_- \cdot \mathbb{E}[X \,|\, \mathcal{G}] = Y \cdot \mathbb{E}[X \,|\, \mathcal{G}]$$

This proves the assertion for every real-valued, integrable and $\mathcal{G}$-measurable random variable Y. Finally suppose that Y is complex valued, $\mathcal{G}$-measurable and integrable. Write $Y = Y_r + i \cdot Y_i$ for real valued Y_r, Y_i random variables. Then Y_r, Y_i are $\mathcal{G}$-measurable and integrable. Hence

$$\mathbb{E}[Y \cdot X \,|\, \mathcal{G}] = \mathbb{E}[Y_r \cdot X \,|\, \mathcal{G}] + i \cdot \mathbb{E}[Y_i \cdot X \,|\, \mathcal{G}] = Y_r \cdot \mathbb{E}[X \,|\, \mathcal{G}] + i \cdot Y_i \cdot \mathbb{E}[X \,|\, \mathcal{G}] = Y \cdot \mathbb{E}[X \,|\, \mathcal{G}]$$

Thus assertion holds for any $\mathcal{G}$-measurable, integrable random variable $Y : \Omega \to \mathbb{C}$. Suppose now that Y is $\mathcal{G}$-measurable and $Y \cdot X$, X are integrable. Define $W_n = \{\omega \in \Omega \,|\, |Y(\omega)| \leq n\}$ and $Y_n = \chi_{W_n} \cdot Y$. Then $\{Y_n\}_{n\in\mathbb{N}}$ is a sequence of integrable $\mathcal{G}$-measurable random variables convergent to Y and $|Y_n \cdot X| \leq |Y \cdot X|$ for every $n \in \mathbb{N}$. Hence

$$Y \cdot \mathbb{E}[X \,|\, \mathcal{G}] = \lim_{n\to+\infty} Y_n \cdot \mathbb{E}[X \,|\, \mathcal{G}] = \lim_{n\to+\infty} \mathbb{E}[Y_n \cdot X \,|\, \mathcal{G}] = \mathbb{E}[Y \cdot X \,|\, \mathcal{G}]$$

and the last equality follow from **(4)** of Theorem 3.1 □

Theorem 3.3 (Tower Property). *Let $\mathcal{G}_2 \subseteq \mathcal{G}_1 \subseteq \mathcal{F}$ be σ-algebras and $X : \Omega \to \mathbb{C}$ be an integrable random variable. Then*

$$\mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{G}_1] \,\big|\, \mathcal{G}_2\big] = \mathbb{E}[X \,|\, \mathcal{G}_2]$$

Proof. Fix $G \in \mathcal{G}_2$. Then also $G \in \mathcal{G}_1$ and

$$\int_G \mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{G}_1] \,\big|\, \mathcal{G}_2\big]\, dP = \int_G \mathbb{E}[X \,|\, \mathcal{G}_1]\, dP = \int_G X\, dP = \int_G \mathbb{E}[X \,|\, \mathcal{G}_2]\, dP$$

Therefore, we derive that $\mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{G}_1] \,\big|\, \mathcal{G}_2\big] = \mathbb{E}[X \,|\, \mathcal{G}_2]$. □

Theorem 3.4. *Let $\mathcal{G}$ be a σ-subalgebra of $\mathcal{F}$, $X : \Omega \to \mathbb{R}$ be an integrable random variable and $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function. Suppose that $\phi(X)$ is integrable. Then*

$$\phi\left(\mathbb{E}[X \,|\, \mathcal{G}]\right) \leq \mathbb{E}[\phi(X) \,|\, \mathcal{G}]$$

Proof. Let L_ϕ be a set of functions $\mathbb{R} \ni x \mapsto a \cdot x + b \in \mathbb{R}$ for $a, b \in \mathbb{R}$ such that $a \cdot x + b \leq \phi(x)$ for every $x \in \mathbb{R}$. Since ϕ is convex, we derive that for every $x \in \mathbb{R}$ we have $\phi(x) = \sup_{l \in L_\phi} l(x)$. Hence

$$\phi\left(\mathbb{E}[X \,|\, \mathcal{G}]\right) = \sup_{l \in L_\phi} l\left(\mathbb{E}[X \,|\, \mathcal{G}]\right) = \sup_{l \in L_\phi} \mathbb{E}[l(X) \,|\, \mathcal{G}] \leq \mathbb{E}[\phi(X) \,|\, \mathcal{G}]$$

□

References

[Monygham, 2018] Monygham (2018). Radon-nikodym theorem, hahn-jordan decomposition and lebesgue decomposition. *github repository: "Monygham/Pedo-mellon-a-minno"*.