Recap on conditional expectations

Here $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\mathcal{A} \subset \mathcal{F}$ is a σ-field. Recall that if $\mathcal{B}$ is a σ-field, $L^1(\Omega, \mathcal{B}, \mathbb{P})$ denotes the set of all real-valued random variables which are $\mathcal{B}$ measurable and integrable (i.e. all measurable $X : (\Omega, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mathbb{E}[|X|] < \infty$).

1 Definitions

$*$ When $X : (\Omega, \mathcal{F}) \to \mathbb{R}$ is a random variable, one defines $\mathbb{E}[X|\mathcal{A}]$ in two cases:

(a) when X is integrable. In this case $\mathbb{E}[X|\mathcal{A}]$ is a random variable X', defined uniquely almost surely (that is, uniquely up to 0 probability events), such that

(1) $X' \in L^1(\Omega, \mathcal{A}, \mathbb{P})$

(2) for every Z real-valued $\mathcal{A}$ measurable bounded random variable, $\mathbb{E}[X'Z] = \mathbb{E}[XZ]$.

The property (2) is called the "characteristic property of conditional expectation". By linearity (2) can be replaced by:

(2') for every Z real-valued $\mathcal{A}$ measurable non-negative bounded random variable, $\mathbb{E}[X'Z] = \mathbb{E}[XZ]$.

By using the Dynkin Lemma (see Exercise 2 in Exercise sheet 8), (2) can be replaced by

(2") for every A belonging to a generating π-system of $\mathcal{A}$, $\mathbb{E}[X'\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A]$.

(b) when X is $[0, \infty]$-valued. In this case $\mathbb{E}[X|\mathcal{A}]$ is a random variable X', defined uniquely almost surely (that is, uniquely up to 0 probability events), such that

(1) $X' \in [0, \infty]$ and X' is $\mathcal{A}$-measurable.

(2) for every Z non-negative random variable, $\mathcal{A}$ measurable, $\mathbb{E}[X'Z] = \mathbb{E}[XZ]$.

The property (2) is called the "characteristic property of conditional expectation". By monotone convergence (2) can be replaced by:

(2') for every Z real-valued $\mathcal{A}$ measurable non-negative bounded random variable, $\mathbb{E}[X'Z] = \mathbb{E}[XZ]$.

By using the Dynkin Lemma, (2) can be replaced by

(2") for every A belonging to a generating π-system of $\mathcal{A}$, $\mathbb{E}[X'\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A]$.

In this case (b), all the expectations are well-defined (with values in $[0, \infty]$) since they only involve random variables with values in $[0, \infty]$.

In practice, to show that $\mathbb{E}[X|\mathcal{A}]$ is some random variable X', one can show that X' satisfies (1) and (2) above (or use some general properties on conditional expectations below).

$*$ When Y is a random variable defined on Ω (with values in any space), one defines

$$\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)],$$

which is a $\sigma(Y)$-measurable random variable. When Y is $\mathbb{R}^n$-valued, by the Doob Dynkin Lemma $\mathbb{E}[X|Y]$ can be written in the form $\phi(Y)$ with $\phi : \mathbb{R}^n \to \mathbb{R}$ measurable.

$*$ When $B \in \mathcal{F}$ is an event, one defines

$$\mathbb{P}(B|\mathcal{A}) = \mathbb{E}[\mathbb{1}_B|\mathcal{A}],$$

which is an $\mathcal{A}$-measurable random variable.

2 Useful properties of conditional expectations

All the random variables are defined on $(\Omega, \mathcal{F}, \mathbb{P})$. As usual, the "almost surely" is implicit in every assertion containing conditional expectations.

(1) If X is integrable or $[0,\infty]$-valued, $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{A}\right]\right] = \mathbb{E}\left[X\right]$;

(2) If X is integrable or $[0,\infty]$-valued and $\mathcal{A}$-measurable, $\mathbb{E}\left[X|\mathcal{A}\right] = X$;

(3) If X is integrable or $[0,\infty]$-valued, $\mathbb{E}\left[X|\{\varnothing,\Omega\}\right] = \mathbb{E}\left[X\right]$;

(4) If X, Y are $[0,\infty]$ valued and $a, b \geq 0$, then

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

If X, Y are integrable and $a, b \in \mathbb{R}$, then

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

(5) If X, Y are integrable or $[0,\infty]$-valued with $X \geq Y$, then $\mathbb{E}\left[X|\mathcal{A}\right] \geq \mathbb{E}\left[Y|\mathcal{A}\right]$.

(6) If X is integrable, $|\mathbb{E}\left[X|\mathcal{A}\right]| \leq \mathbb{E}\left[|X||\mathcal{A}\right]$.

(7) If X is integrable or $[0,\infty]$-valued, and Y is independent of X then $\mathbb{E}\left[X|Y\right] = \mathbb{E}\left[X\right]$.

We have the following limit theorems.

(8) (conditional monotone convergence) If (X_n) is a weakly increasing sequence of $[0,\infty]$-valued random variables and $X = \lim \uparrow X_n$, then

$$\mathbb{E}\left[X|\mathcal{A}\right] = \lim_{n\to\infty} \mathbb{E}\left[X_n|\mathcal{A}\right].$$

(9) (conditional Fatou's lemma) If (X_n) is a sequence of $[0,\infty]$-valued random variables, then

$$\mathbb{E}\left[\liminf_{n\to\infty} X_n|\mathcal{A}\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[X_n|\mathcal{A}\right].$$

(10) (conditional dominated convergence) Let (X_n) be a sequence of integrable random variables converging almost surely to X. Assume that there is a random variable Z such that for every $n \geq 1$, almost surely $|X_n| \leq Z$ and $\mathbb{E}\left[Z\right] < \infty$. Then

$$\mathbb{E}\left[X|\mathcal{A}\right] = \lim_{n\to\infty} \mathbb{E}\left[X_n|\mathcal{A}\right] \qquad \text{a.s. and in } L^1.$$

(11) (conditional Jensen's inequality) Let f is a nonnegative convex function. If X is integrable, then

$$f\left(\mathbb{E}\left[X|\mathcal{A}\right]\right) \leq \mathbb{E}\left[f(X)|\mathcal{A}\right].$$

By taking $\mathcal{A} = \{\varnothing, \Omega\}$ we get $f(\mathbb{E}\left[X\right]) \leq \mathbb{E}\left[f(X)\right]$ (Jensen's inequality).

3 Further properties of conditional expectations

The following properties are useful when several random variables and/or σ-fields are present.

(12) (factorizing out) Let X, Y be random variables with either $X, Y \in [0,\infty]$, or X and XY integrable. Assume that Y is $\mathcal{A}$-measurable. Then

$$\mathbb{E}\left[YX|\mathcal{A}\right] = Y\mathbb{E}\left[X|\mathcal{A}\right].$$

(13) (tower property) Let $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{F}$ be σ-fields. Then for X integrable or $[0,\infty]$-valued,

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{A}_2\right]|\mathcal{A}_1\right] = \mathbb{E}\left[X|\mathcal{A}_1\right].$$

(14) (adding independent information) Let $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{F}$ be σ-fields. Assume that X is integrable or $[0,\infty]$-valued. If $\mathcal{A}_2$ is independent of $\sigma(\sigma(X), \mathcal{A}_1)$, then

$$\mathbb{E}\left[X|\sigma(\mathcal{A}_1, \mathcal{A}_2)\right] = \mathbb{E}\left[X|\mathcal{A}_1\right].$$