35. Espérance conditionnelle [BL07, §VI.2, p156–157]

ÉNONCÉ

PROPOSITION. [ESPÉRANCE CONDITIONNELLE]
Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et soit $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. Soit de plus X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$, intégrable ($X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$). Alors il existe une unique (presque sûrement) variable aléatoire Z telle que :

(i) Z est $\mathcal{B}$-mesurable,
(ii) pour tout $B \in \mathcal{B}$, $\mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[Z\mathbb{1}_B]$.

De plus, Z est intégrable.

DÉFINITION. [ESPÉRANCE CONDITIONNELLE]
Z est appelée espérance conditionnelle de X sachant $\mathcal{B}$, et est notée $\mathbb{E}[X \mid \mathcal{B}]$.

DÉVELOPPEMENT

1. Montrons l'unicité.
   Soient Z_1 et Z_2 deux variables satisfaisants (i) et (ii).
   Z_1 et Z_2 sont $\mathcal{B}$-mesurables donc $W = \mathbb{1}_{Z_1 > Z_2}$ aussi et alors $\mathbb{E}[Z_1 W] = \mathbb{E}[XW] = \mathbb{E}[Z_2 W]$. D'où $\mathbb{E}[(Z_1 - Z_2)\mathbb{1}_{Z_1 > Z_2}] = 0$.
   Comme par ailleurs $(Z_1 - Z_2)\mathbb{1}_{Z_1 > Z_2} \geq 0$ p.s., on en déduit $(Z_1 - Z_2)\mathbb{1}_{Z_1 > Z_2} = 0$ p.s..
   Ou encore $Z_1 \leq Z_2$ p.s.. Par symétrie on a $Z_1 = Z_2$ p.s..

2. Montrons l'existence. On procède en plusieurs étapes :

   - Supposons $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$:
     Rappelons que $L^2(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de HILBERT. $\mathbb{L}^2(\Omega, \mathcal{B}, \mathbb{P})$ en est un sous-espace vectoriel convexe fermé (car une limite de fonctions $\mathcal{B}$-mesurables est $\mathcal{B}$-mesurable), on peut donc considérer la projection Z de X sur $\mathbb{L}^2(\Omega, \mathcal{B}, \mathbb{P})$. Alors Z satisfait (i) et comme pour tout $B \in \mathcal{B}$, on a $\mathbb{1}_B \in \mathbb{L}^2(\Omega, \mathcal{B}, \mathbb{P})$, on a :

     $$\forall B \in \mathcal{B}, \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[(X - Z)\mathbb{1}_B] + \mathbb{E}[Z\mathbb{1}_B] = \mathbb{E}[Z\mathbb{1}_B]$$

     Donc Z satisfait (ii). Ainsi $Z = \mathbb{E}[X \mid \mathcal{B}]$ et Z est bien intégrable car L^2.

     NOTE.
     - Si $X \geq 0$ p.s., alors $\mathbb{E}[X \mid \mathcal{B}] \geq 0$ p.s.. En effet, prenant $B = \{\mathbb{E}[X \mid \mathcal{B}] < 0\} \in \mathcal{B}$, on obtient dans (ii) que $\mathbb{1}_B = 0$ p.s..
     - Par linéarité de la projection, on a que si $X_1, X_2 \in L^2$, alors $\mathbb{E}[X_1 + X_2 \mid \mathcal{B}] = \mathbb{E}[X_1 \mid \mathcal{B}] + \mathbb{E}[X_2 \mid \mathcal{B}]$.
     - prenant $B = \Omega$, on obtient $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{B}]]$.

   - Supposons $X \geq 0$ p.s. :
     Pour $n \in \mathbb{N}$, on pose $X_n = \min(X, n) \in L^2$.
     Par le point précédent, les $(\mathbb{E}[X_n \mid \mathcal{B}])_{n \in \mathbb{N}}$ existent et la suite est croissante puisque $(X_n)_{n \in \mathbb{N}}$ l'est. Soit alors $Z = \lim_{n \to +\infty} \uparrow \mathbb{E}[X_n \mid \mathcal{B}]$. Z est $\mathcal{B}$-mesurable et pour tout $B \in \mathcal{B}$, on a :

     $$\begin{aligned} \mathbb{E}[X\mathbb{1}_B] &= \lim_{n \to +\infty} \uparrow \mathbb{E}[X_n \mathbb{1}_B] && \text{par convergence monotone} \\ &= \lim_{n \to +\infty} \uparrow \mathbb{E}[\mathbb{E}[X_n \mid \mathcal{B}]\mathbb{1}_B] && \text{par (ii) vraie dans le cas } L^2 \\ &= \mathbb{E}[Z\mathbb{1}_B] && \text{par convergence monotone} \end{aligned}$$

     Donc $Z = \mathbb{E}[X \mid \mathcal{B}]$ et en prenant $B = \Omega$, on a à nouveau $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{B}]]$.
     De plus $\mathbb{E}[X \mid \mathcal{B}] \geq 0$ p.s..

   - Enfin, si $X \in L^1$:
     On décompose X sous la forme $X = X^+ - X^-$ et on pose $Z = \mathbb{E}[X^+ \mid \mathcal{B}] - \mathbb{E}[X^- \mid \mathcal{B}]$, qui est $\mathcal{B}$-mesurable. On a que Z est intégrable puisque :

     $$\mathbb{E}[|Z|] = \mathbb{E}[\mathbb{E}[X^+ \mid \mathcal{B}]] + \mathbb{E}[\mathbb{E}[X^- \mid \mathcal{B}]] = \mathbb{E}[X^+] + \mathbb{E}[X^-] = \mathbb{E}[|X|]$$

     On remarque alors que $\mathbb{E}[Z] = \mathbb{E}[X]$.
     De plus, on a pour $B \in \mathcal{B}$:

     $$\begin{aligned} \mathbb{E}[X\mathbb{1}_B] &= \mathbb{E}[(X^+ - X^-)\mathbb{1}_B)] \\ &= \mathbb{E}[X^+\mathbb{1}_B] - \mathbb{E}[X^-\mathbb{1}_B] \\ &= \mathbb{E}[\mathbb{E}[X^+ \mid \mathcal{B}]\mathbb{1}_B] - \mathbb{E}[\mathbb{E}[X^- \mid \mathcal{B}]\mathbb{1}_B] \\ &= \mathbb{E}[(\mathbb{E}[X^+ \mid \mathcal{B}] - \mathbb{E}[X^- \mid \mathcal{B}])\mathbb{1}_B] \\ &= \mathbb{E}[Z\mathbb{1}_B] \end{aligned}$$

     Donc $Z = \mathbb{E}[X \mid \mathcal{B}]$.

REMARQUE. On a donc montré de plus que $\mathbb{E}[\mathbb{E}[X \mid \mathcal{B}]] = \mathbb{E}[X]$.