Chapitre 2 : Espérance conditionnellement à une tribu

2.1 Introduction : notions d'espérance conditionnelle déjà rencontrées

Dans ce qui suit $(\Omega, \mathcal{F}, \mathbb{P})$ est donné.

2.1.1 Espérance conditionnellement à un évènement

Soit $A \in \mathcal{F}$ avec $\mathbb{P}(A) > 0$. On note $\mathbb{P}(\cdot \,|\, A)$ la mesure de définie sur $(\Omega, \mathcal{F})$ par

$$\forall \Gamma \in \mathcal{F}, \quad \mathbb{P}(\Gamma \,|\, A) = \frac{\mathbb{P}(\Gamma \cap A)}{\mathbb{P}(A)}.$$

C'est une mesure de probabilités sur $(\Omega, \mathcal{F})$. Elle vérifie $\mathbb{P}(A|A) = 1$ et $\mathbb{P}(\Gamma|A) = 0$ pour tout $\Gamma \in \mathcal{F}$ t.q. $\Gamma \cap A = \emptyset$. On l'appelle la mesure de probabilités sachant A.

Soit X v.a. On définit l'espérance de X sachant l'évènement A comme

$$\mathbb{E}_A(X) = \mathbb{E}(X|A) := \int_\Omega X(\omega)\mathbb{P}(d\omega \,|\, A) = \frac{\mathbb{E}(\mathbf{1}_A X)}{\mathbb{P}(A)} = \frac{\int_A X(\omega)\mathbb{P}(d\omega)}{\mathbb{P}(A)}.$$

2.1.2 Espérance d'une v.a. conditionnellement à une autre : cas discret

Soient X et Y deux v.a. discrètes (à val. resp. dans E et E'), avec en particulier $\mathbb{P}(Y = y) > 0$ pour tout $y \in E'$.

On souhaite définir l'espérance de X sachant Y. On définit

$$\mathbb{E}(X|Y = y) := \mathbb{E}(X \,|\, \{Y = y\}) = \frac{\mathbb{E}[X\mathbf{1}_{\{Y=y\}}]}{\mathbb{P}(Y = y)}, \forall y \in E'$$

(ce qui est possible car $\mathbb{P}(Y = y) > 0$).

On note ensuite $\varphi(y) = \mathbb{E}(X|Y = y)$, $y \in E'$ et on définit

$$\mathbb{E}(X|Y) := \varphi(Y).$$

L'objet $\mathbb{E}(X|Y)$ est une v.a. $\sigma(Y)$-mesurable (car Y est $\sigma(Y)$-mes. et $\varphi : E' \to \mathbb{R}$ est mes.).

2.1.3 Espérance d'une v.a. conditionnellement à une autre : cas à densité bivariée

Soit (X, Y) à valeurs dans $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ et de loi à densité bivariée $f_{X,Y}(x, y)$, i.e.

$$\forall B \times C \in \mathcal{B}(\mathbb{R}^2), \quad \mathbb{P}(X \in B; Y \in C) = \int\int_{B \times C} f_{X,Y}(x, y)\, dxdy.$$

Le loi marginale $\mathbb{P}_Y(dy)$ est donnée par

$$\mathbb{P}_Y(dy) = f_Y(y)dy \quad \text{avec} \quad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\, dx$$

(**En effet :** [Justif. par Fubini au tableau]).

Ainsi la loi de Y est à densité par rapport à la mesure de Lebesgue, donc $\mathbb{P}(Y = y) = 0$, $\forall y \in \mathbb{R}$.

☛ On ne peut pas définir $\mathbb{E}(X|Y=y)$ comme précédemment.

On pose

$$\mathbb{E}(X|Y=y) := \int_{\mathbb{R}} x \frac{f_{X,Y}(x,y)}{f_Y(y)} dx, \quad \forall y \in \mathbb{R}.$$

Puis à nouveau $\varphi(y) = \mathbb{E}(X|Y=y)$, $y \in \mathbb{R}$ et

$$\mathbb{E}(X|Y) := \varphi(Y).$$

Résultat admis : $\varphi : \mathbb{R} \to \mathbb{R}$ est mesurable (cf Remarque 2.1.1 du poly).

A nouveau, comme φ est mesurable la v.a. $\mathbb{E}(X|Y)$ est $\sigma(Y)$-mesurable.

Question : Y a-t-il un caractère commun à $\mathbb{E}(X|Y)$ défini ici et en sous-section 2.1.2 ?

2.1.4 Vers une caractérisation commune des espérances conditionnelles vues précédemment

Considérons la v.a. $\mathbb{E}(X|Y)$ définie à la sous-section 2.1.3. Soit $A \in \sigma(Y)$ on a $A = \{Y \in B\}$ pour un certain $B \in \mathcal{B}(\mathbb{R})$ (cf Feuille de TD 2 Exercice 3). On a

$$\mathbb{E}\big[\mathbf{1}_A \mathbb{E}(X|Y)\big] = \mathbb{E}\big[\mathbf{1}_A X\big].$$

Calcul : [au tableau]

En résumé la v.a. $\mathbb{E}(X|Y)$ vérifie :
i) Elle est $\sigma(Y)$-mesurable.
ii) Pour tout A dans $\sigma(Y)$ on a $\mathbb{E}\big[\mathbf{1}_A \mathbb{E}(X|Y)\big] = \mathbb{E}\big[\mathbf{1}_A X\big]$.

Exercice 2.1.1 : Montrer qu'il en est de même dans le cas discret rappelé en sous-section 2.1.2.

En fait c'est comme ça (par les points i) et ii)) qu'on définira plus loin $\mathbb{E}(X|Y)$ (cf sous-section 2.2.3), en utilisant une nouvelle notion :

2.2 Espérance conditionnellement à une tribu

On a $\mathcal{H} \subset \mathcal{F}$ donnée. Pour $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ on cherche à définir une v.a. $\mathbb{E}(X|\mathcal{H})$ qui vérifie :

i) $\mathbb{E}(X|\mathcal{H})$ est $\mathcal{H}$-mesurable

ii) Pour tout $A \in \mathcal{H}$, on a $\mathbb{E}[\mathbf{1}_A \mathbb{E}(X|\mathcal{H})] = \mathbb{E}[\mathbf{1}_A X]$.

2.2.1 Première approche : le cadre L^2

On suppose que $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $\mathbb{E}|X|^2 < +\infty$.
On note $L^2(\Omega, \mathcal{H}, \mathbb{P})$ l'ensemble des variables aléatoires $\mathcal{H}$-mesurables et de carré intégrable. On rappelle les résultats suivants :

Lemme (2.2.1)

i) L'espace $L^2(\Omega, \mathcal{F}, \mathbb{P})$ muni du produit scalaire $(X, Y) \mapsto \mathbb{E}(XY)$ est un espace de Hilbert.
ii) Le sous-espace $L^2(\Omega, \mathcal{H}, \mathbb{P}) \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ est une sous-espace fermé de $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Pour $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ on considère son projeté orthogonal sur $L^2(\Omega, \mathcal{H}, \mathbb{P})$.
☛ On décide de le noter $\mathbb{E}(X|\mathcal{H})$.

La v.a. $\mathbb{E}(X|\mathcal{H})$ ainsi définie vérifie les points i) et ii) (du transparent précédent).
[Explication et dessin, au tableau]

Remarque : On a $||\mathbb{E}(X|\mathcal{H})||_{L^2} \leq ||X||_{L^2}$. En effet

$$||X||_{L^2}^2 = ||X - \mathbb{E}(X|\mathcal{H}) + \mathbb{E}(X|\mathcal{H})||_{L^2}^2 = ||X - \mathbb{E}(X|\mathcal{H})||_{L^2}^2 + ||\mathbb{E}(X|\mathcal{H})||_{L^2}^2$$

et $||X - \mathbb{E}(X|\mathcal{H})||_{L^2}^2 \geq 0$.

2.2.2 Extension au cadre L^1 et premières propriétés

Si $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ on a le résultat suivant.

Théorème-Définition (2.2.1)

Soit X un vecteur aléatoire, supposé intégrable, i.e. $\mathbb{E}|X| < +\infty$.
Alors il existe une v.a. Y vérifiant
1) Y est $\mathcal{H}$-mesurable
2) Pour tout $A \in \mathcal{H}$, on a $\mathbb{E}(\mathbf{1}_A Y) = \mathbb{E}(\mathbf{1}_A X)$.
Cette v.a. Y est unique à l'égalité p.s. près et se note $\mathbb{E}(X|\mathcal{H})$.
Elle est appelée espérance de X sachant $\mathcal{H}$ ou espérance de X conditionnellement à $\mathcal{H}$.

Preuve : Existence : cf poly.
Unicité : cf Proposition 2.2.2 ci-après.

Remarque (Proposition 2.2.1) : Dans le théorème-définition 2.2.1 la condition 2) peut être remplacée de façon équivalente par
2') Pour toute v.a.r. Z bornée et $\mathcal{H}$-mesurable on a $\mathbb{E}(ZY) = \mathbb{E}(ZX)$.

Proposition (2.2.2)

Si Y et Y' vérifient les points 1) et 2) du théorème-définition 2.2.2 on a $Y = Y'$ p.s., d'où l'égalité p.s. près.

Preuve : [au tableau]

Lemme (2.2.2)

Pour $X \in L^1$ on a $\mathbb{E}(X|\mathcal{H}) \in L^1$.

Preuve : cf Feuille TD 3, exercice 2, point 1).

On liste maintenant des propriétés qui reviendront continuellement dans les calculs d'espérance conditionnelle.

Certains seront démontrés en TD (Feuille 3, exercice 1). Pour les autres cf poly.

Propriété (2.2.1)

On a (les égalités ou inégalités sont au sens p.s.) :
i) Pour $a, b \in \mathbb{R}$, $\mathbb{E}(aX + bY|\mathcal{H}) = a\mathbb{E}(X|\mathcal{H}) + b\mathbb{E}(Y|\mathcal{H})$ *(linéarité de l'espérance conditionnelle).*
ii) Si $X \geq 0$ *alors* $\mathbb{E}(X|\mathcal{H}) \geq 0$.
iii) $\mathbb{E}[\mathbb{E}(X|\mathcal{H})] = \mathbb{E}(X)$.
iv) Si X *est* $\mathcal{H}$*-mesurable alors* $\mathbb{E}(X|\mathcal{H}) = X$.
v) Si Y *est* $\mathcal{H}$*-mesurable (mais* X *v.a. intégrable quelconque) alors* $\mathbb{E}(YX|\mathcal{H}) = Y\mathbb{E}(X|\mathcal{H})$.
vi) Si X *est indépendante de* $\mathcal{H}$ *(i.e.* X *ind. de* $\mathbf{1}_A$, $\forall A \in \mathcal{H}$*), alors* $\mathbb{E}(X|\mathcal{H}) = \mathbb{E}(X)$.
vii) Si $\mathcal{G}$ *est une sous-tribu avec* $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$*, alors*

$$\mathbb{E}\big[\mathbb{E}(X|\mathcal{G})\,|\,\mathcal{H}\big] = \mathbb{E}(X|\mathcal{H}) = \mathbb{E}\big[\mathbb{E}(X|\mathcal{H})\,|\,\mathcal{G}\big].$$

Définition 2.2.1 : Pour $A \in \mathcal{F}$ on note $\mathbb{P}(A|\mathcal{H}) = \mathbb{E}(\mathbf{1}_A|\mathcal{H})$.

2.2.3 Espérance d'une v.a. conditionnellement à une autre

Definition (2.2.2)

Soient X et Y définies sur $(\Omega, \mathcal{F}, \mathbb{P})$, avec X intégrable. On note

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

Lemme (2.2.3, Lemme de Doob)

Soit Y v.a. Une variable aléatoire Z est $\sigma(Y)$-mesurable si et seulement si il existe φ borélienne t.q. $Z = \varphi(Y)$.

En conséquence il existe φ borélienne t.q.

$$\mathbb{E}(X|Y) = \varphi(Y).$$

Cela nous amène à poser la notation

$$\mathbb{E}(X|Y = y) := \varphi(y) \qquad (2.2.3)$$

Question : Avec cette nouvelle définition retrouve-t-on les cas à densité bivariée et discret vus en section 2.1 ?

Cas à densité bivariée. La v.a. $\mathbb{E}(X|Y)$ définie dans la sous-section 2.1.3 vérifie :
i) Elle est $\sigma(Y)$-mesurable.
ii) Pour tout A dans $\sigma(Y)$ on a $\mathbb{E}\big[\mathbf{1}_A \mathbb{E}(X|Y)\big] = \mathbb{E}\big[\mathbf{1}_A X\big]$.

Donc par unicité c'est $\mathbb{E}[X|\sigma(Y)]$, c'est à dire $\mathbb{E}(X|Y)$ *au sens de la nouvelle définition*.
C'est donc que $\mathbb{E}(X|Y)$ au sens de la nouvelle définition vaut $\varphi(Y)$ avec la fonction $\varphi(y)$ définie dans la sous-section 2.1.3
C'est donc que pour tout $y \in \mathbb{R}$ la quantité $\mathbb{E}(\, X \mid Y = y)$ *au sens de la nouvelle définition* vaut

$$\varphi(y) = \int_{\mathbb{R}} x \, \frac{f_{X,Y}(x,y)}{f_Y(y)} dx.$$

La nouvelle définition aboutit donc à la même valeur pour $\mathbb{E}(\, X \mid Y = y)$ que l'ancienne.

Cas discret : Soit $y \in E'$; en partant de la nouvelle définition on retrouve

$$\mathbb{E}(\, X \,|\, Y = y) = \frac{\mathbb{E}[X\mathbf{1}_{\{Y=y\}}]}{\mathbb{P}(Y = y)},$$

ce qui est cohérent avec l'ancienne.

[Calcul au tableau]

2.2.4 Autres propriétés

Proposition (2.2.4)

Soient X et Y définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. Supposons que X est $\mathcal{H}$-mesurable et que Y est indépendante de $\mathcal{H}$ (donc de X).
Alors pour tout fonction f borélienne telle que $f(X,Y) \in L^1$ on a

$$\mathbb{E}[f(X,Y) \,|\, \mathcal{H}] = F(X) \quad \text{où} \quad F(x) = \mathbb{E}[f(x,Y)], \ \forall x \in \mathbb{R}.$$

Remarque 2.2.2 : [au tableau]

Grenoble INP Ensimag

Théorème (2.2.1)

1) (Convergence monotone) : Si (X_n) *est une suite de v.a. positives qui tend en croissant vers* $X \in L^1$ *alors*

$$\mathbb{E}(X_n|\mathcal{H}) \xrightarrow[n\to\infty]{\text{p.s.}} \mathbb{E}(X|\mathcal{H}).$$

2) (Fatou) : Si $X_n \geq Z$ *pour tout* n *avec* $Z \in L^1$ *et* $\mathbb{E}|X_n| < \infty$ *pour tout* n *et* $\mathbb{E}|\liminf_n X_n| < \infty$ *alors*

$$\mathbb{E}(\liminf_n X_n \,|\, \mathcal{H}) \leq \liminf_n \mathbb{E}(X_n \,|\, \mathcal{H}) \quad \text{p.s.}$$

3) (Convergence dominée) : Si

$$X_n \xrightarrow[n\to\infty]{\text{p.s.}} X \in L^1$$

et s'il existe $Z \in L^1$ *t.q.* $|X_n| \leq Z$ *p.s. pour tout* n *alors,*

$$\mathbb{E}(X_n|\mathcal{H}) \xrightarrow[n\to\infty]{\text{p.s.}} \mathbb{E}(X|\mathcal{H}).$$

4) (Jensen) : Si $X \in L^1$ *et* $\varphi : \mathbb{R} \to \mathbb{R}$ *est mesurable convexe avec* $\varphi(X) \in L^1$ *alors*

$$\varphi\big(\mathbb{E}(X|\mathcal{H})\big) \leq \mathbb{E}(\varphi(X) \,|\, \mathcal{H}) \quad \text{p.s.}$$