Probabilités et Intégration
Licence 3, S6 Option Mathématiques Appliquées
et DUAS 1, S2

Table des matières

Chapitre 1

Espérance conditionnelle

1.1 Définition et propriétés de l'espérance conditionnelle

Dans tout ce paragraphe, on se place sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Définition 1 *i) On dit que $\mathcal{A} \subset \mathcal{P}(\Omega)$ est une sous-tribu de $\mathcal{F}$ si $\mathcal{A}$ est une tribu telle que $\mathcal{A} \subset \mathcal{F}$.*
ii) On dit que la fonction $f : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est $\mathcal{A}$-mesurable si pour tout $B \in \mathcal{B}(\mathbb{R})$, $f^{-1}(B) \in \mathcal{A}$.

Définition 2 *Soit X une variable aléatoire intégrable. L'espérance conditionnelle de X sachant $\mathcal{A}$ est une variable aléatoire $\mathbb{E}(X|\mathcal{A}) : \Omega \to \mathbb{R}$ telle que :*
1) $\mathbb{E}(X|\mathcal{A})$ est $\mathcal{A}$-mesurable,
2) pour tout $A \in \mathcal{A}$,

$$\int_A X d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{A}) d\mathbb{P}.$$

Proposition 1 *L'espérance conditionnelle de X sachant $\mathcal{A}$ existe et est unique (modulo l'égalité presque-sûre).*

Proposition 2 *Dans toute la suite, X est une variable aléatoire $\mathcal{F}$-mesurable et intégrable.*
0) L'espérance conditionnelle est linéaire et croissante. Autrement dit, soit Y une variable aléatoire $\mathcal{F}$-mesurable et intégrable et $(\lambda_1, \lambda_2) \in \mathbb{R}^2$,

$$\mathbb{E}(\lambda_1 X + \lambda_2 Y|\mathcal{A}) = \lambda_1 \mathbb{E}(X|\mathcal{A}) + \lambda_2 \mathbb{E}(Y|\mathcal{A}) p.s.,$$

et si $X \leq Y$ p.s., $\mathbb{E}(X|\mathcal{A}) \leq \mathbb{E}(Y|\mathcal{A})$ p.s.
1) $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\mathcal{A}))$.
2) Si X est $\mathcal{A}$ mesurable, $\mathbb{E}(X|\mathcal{A}) = X$ presque sûrement.

3) Soit $(A_n)_{n\in\mathbb{N}}$ une partition de Ω avec $A_n \in \mathcal{F}$ et $\mathbb{P}(A_n) > 0$. Si $\mathcal{A}$ est la tribu engendrée par la partition $(A_n)_{n\in\mathbb{N}}$ alors

$$\mathbb{E}(X|\mathcal{A}) = \sum_{n\in\mathbb{N}} \mathbb{E}(X \mid A_n)\mathbb{I}_{A_n}.$$

4) Si $\mathcal{A}$ est la tribu triviale $\{\Omega, \emptyset\}$ alors $\mathbb{E}(X|\mathcal{A}) = \mathbb{E}(X)$.

5) Soit Z une variable aléatoire $\mathcal{A}$-mesurable telle que ZX soit intégrable alors $\mathbb{E}(ZX|\mathcal{A}) = Z\mathbb{E}(X|\mathcal{A})$ presque-sûrement.

6) Soient $\mathcal{A}_2 \subset \mathcal{A}_1$ des sous tribus de $\mathcal{F}$. On a

$$\mathbb{E}\{\mathbb{E}(X|\mathcal{A}_1)|\mathcal{A}_2\} = \mathbb{E}\{\mathbb{E}(X|\mathcal{A}_2)|\mathcal{A}_1\} = \mathbb{E}(X|\mathcal{A}_2).$$

Intéressons à présent au cas où la tribu $\sigma(X) = \{X^{-1}(B);\ B \in \mathcal{B}(\mathbb{R})\}$ est indépendante de la sous tribu $\mathcal{A}$.

Proposition 3 *Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une variable aléatoire intégrable. Si la sous tribu $\mathcal{A}$ est indépendante de $\sigma(X)$ alors, $\mathbb{E}(X|\mathcal{A}) = \mathbb{E}(X)$.*

1.2 Espérance conditionnelle pour des variables aléatoires de carré intégrable

Proposition 4 *Soit X une variable aléatoire appartenant à l'espace $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Alors, $\mathbb{E}(X|\mathcal{A})$ est la projection orthogonale de X sur le sous espace $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Autrement dit,*

$$\|X - \mathbb{E}(X|\mathcal{A})\|_2 = \inf\{\|X - Y\|_2,\ Y \in L^2(\Omega, \mathcal{A}, \mathbb{P})\}.$$

1.3 Espérance conditionnelle par rapport à une variable aléatoire

Définition 3 *Soient $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ et $Y : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ deux variables aléatoires. L'espérance conditionnelle de X sachant Y (notée $\mathbb{E}(X|Y)$) est l'espérance conditionnelle $\mathbb{E}(X|\sigma(Y))$.*

Le résultat important suivant montre que $\mathbb{E}(X|Y)$ est une fonction mesurable de Y.

Lemme 1 (Lemme de Doob) *Soient $Y : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ et $Z : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ deux variables aléatoires mesurables. Si Z est $\sigma(Y)$-mesurable alors il existe une fonction $h : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable telle que $Z = h(Y)$.*

En conséquence $\mathbb{E}(X|Y) = h(Y)$. Ainsi la notation $\mathbb{E}(X|Y = y)$ a un sens même si $\mathbb{P}(Y = y) = 0$ (elle doit être comprise comme la valeur au point y de la fonction $h(\cdot) := \mathbb{E}(X|Y = \cdot)$.

1.3.1 Cas où Y est une variable aléatoire discrète

Soit $Y : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{P}(E))$ où $E = \{x_1, x_2, \ldots\}$ est un espace au plus dénombrable. D'après le point 3) de la Proposition 2,

$$\mathbb{E}(X|Y) = \sum_{n\geq 1} \mathbb{E}[X \mid Y^{-1}(\{x_n\})] \mathbb{I}_{Y^{-1}(\{x_n\})}.$$

En utilisant la notation moins rigoureuse $Y^{-1}(\{x_n\}) = \{Y = x_n\}$, on a

$$\mathbb{E}(X|Y) = \sum_{n\geq 1} \mathbb{E}(X \mid Y = x_n) \mathbb{I}_{\{Y=x_n\}}.$$

1.3.2 Cas où la loi de Y admet une densité par rapport à la mesure de Lebesgue

Soit $(X, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ un couple de variables aléatoires. On suppose dans ce paragraphe que la loi de (X, Y) admet une densité $f_{(X,Y)}$ par rapport à la mesure de Lebesgue.
On note $f_Y(\cdot)$ la densité de la variable aléatoire Y.

Définition 4 *La densité conditionnelle de X sachant $Y = y$ est la fonction mesurable positive définie par :*

$$f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}.$$

Proposition 5 *Pour toute fonction $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable et bornée, $\mathbb{E}(g(X)|Y) = h(Y)$, où $h : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une fonction mesurable définie par*

$$h(y) = \int_{\mathbb{R}} g(x) f_{X|Y=y}(x) dx.$$

Chapitre 2

Convergence de variables aléatoires

2.1 Convergence en probabilité

2.1.1 Définition et propriétés

Définition 5 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires et* X *une variable aléatoire, toutes ces variables étant définies sur le même espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$ *et à valeurs dans* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*. On dit que* X_n *converge en probabilité vers* X *et on note* $X_n \xrightarrow{\mathbb{P}} X$ *si pour tout* $\varepsilon > 0$, $\mathbb{P}(|X_n - X| \geq \varepsilon) \to 0$ *lorsque* n *tend vers l'infini.*

Définition 6 *Soit* $X_n = (X_{n,1}, \ldots, X_{n,p})^\top$ *et* $X = (X_1, \ldots, X_p)^\top$ *des vecteurs aléatoires. On dit que* X_n *converge en probabilité vers* X *si pour tout* $i = 1, \ldots, p$, $X_{n,i}$ *converge en probabilité vers* X_i*.*

Proposition 6 *La suite de vecteurs aléatoires* X_n *converge en probabilité vers* X *si et seulement si pour tout* $\varepsilon > 0$, $\mathbb{P}(\|X_n - X\| > \varepsilon) \to 0$ *lorsque* n *tend vers l'infini. Ici la norme* $\|\cdot\|$ *est une norme quelconque sur* $\mathbb{R}^p$*, toutes les normes étant équivalentes.*

On considère la distance définie pour tout $(X, Y) \in L^0(\Omega, \mathcal{F}, \mathbb{P}) \times L^0(\Omega, \mathcal{F}, \mathbb{P})$ par

$$d(X, Y) = \mathbb{E}\left(\frac{|X - Y|}{1 + |X - Y|}\right).$$

Proposition 7 *Une suite* $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité vers* X *si et seulement si* $d(X_n, X) \to 0$ *lorsque* n *tend vers l'infini.*

Proposition 8 *Si* $(X_n)_{n\in\mathbb{N}}$ *converge en probabilité alors sa limite est unique presque-sûrement.*

Proposition 9 *Soient* (X_n) *et* X *des variables aléatoires définies sur le même espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$ *et à valeurs dans* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$. *Si* $f : \mathbb{R}^p \to \mathbb{R}$ *est une fonction continue et si* $X_n \xrightarrow{\mathbb{P}} X$ *alors* $f(X_n) \xrightarrow{\mathbb{P}} f(X)$.

2.1.2 Loi faible des grands nombres

Soit $(X_k)_{k\in\mathbb{N}}$ une suite de variables aléatoires réelles. On s'intéresse dans ce paragraphe à la convergence en probabilité de la somme partielle $S_n := X_1 + \ldots + X_n$.

Proposition 10 *Soit* $(X_k)_{k\in\mathbb{N}}$ *une suite de variables aléatoires réelles que l'on suppose de même loi, indépendantes et intégrables. Alors* $S_n/n \xrightarrow{\mathbb{P}} \mathbb{E}(X_1)$.

2.2 Convergence presque-sûre

2.2.1 Définitions

Définition 7 *Soient* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires et* X *une variable aléatoire. On dit que* X_n *converge presque-sûrement vers* X *(et on note* $X_n \xrightarrow{ps} X$*) si*

$$\mathbb{P}(\{X_n \to X\}) = \mathbb{P}(\{\omega;\ X_n(\omega) \to X(\omega)\}) = 1.$$

Définition 8 *Soit* $(A_n)_{n\geq 1}$ *une suite d'éléments de* $\mathcal{F}$.

i) On appelle limite supérieure de la suite $(A_n)_{n\geq 1}$ *l'événement* $\overline{\lim} A_n$ *réalisé si une infinité de* A_n *sont réalisés. Autrement dit* $\omega \in \overline{\lim} A_n$ *si pour tout* $k \geq 1$, *il existe* $n \geq k$ *tel que* $\omega \in A_n$. *Ainsi,*

$$\overline{\lim} A_n = \bigcap_{k\geq 1} \bigcup_{n\geq k} A_n.$$

i) On appelle limite inférieure de la suite $(A_n)_{n\geq 1}$ *l'événement* $\underline{\lim} A_n$ *réalisé si tous les* A_n *sont réalisés à partir d'un certain rang. Autrement dit* $\omega \in \underline{\lim} A_n$ *s'il existe* $k \geq 1$ *tel que pour tout* $n \geq k$, $\omega \in A_n$. *Ainsi,*

$$\underline{\lim} A_n = \bigcup_{k\geq 1} \bigcap_{n\geq k} A_n.$$

2.2.2 Lemme de Borel-Cantelli

Proposition 11 (Lemme de Borel Cantelli) *Soit* (A_n) *une suite d'éléments de* $\mathcal{F}$.

i)

$$\sum_{n\geq 1} \mathbb{P}(A_n) < \infty \Rightarrow \mathbb{P}\left(\overline{\lim} A_n\right) = 0.$$

ii) Si les événements $(A_n)_{n\geq 1}$ sont indépendants,

$$\sum_{n\geq 1} \mathbb{P}(A_n) = \infty \Rightarrow \mathbb{P}\left(\overline{\lim} A_n\right) = 1.$$

Corollaire 1 *Si pour tout $\varepsilon > 0$,*

$$\sum_{n\geq 1} \mathbb{P}(|X_n - X| \geq \varepsilon) < \infty,$$

alors $X_n \xrightarrow{ps} X$.

Proposition 12 *La convergence presque-sûre est conservée par les applications continues. Autrement dit, si f est une fonction continue et si $X_n \xrightarrow{ps} X$ alors $f(X_n) \xrightarrow{ps} f(X)$.*

2.2.3 Convergence en probabilité et convergence presque-sûre

La convergence presque-sûre est plus forte que la convergence en probabilité.

Proposition 13 *Si $X_n \xrightarrow{ps} X$ alors $X_n \xrightarrow{\mathbb{P}} X$.*

Proposition 14 *La suite X_n converge en probabilité vers X si et seulement si de toute suite extraite de $(X_n)_{n\geq 1}$ on peut extraire une sous suite qui converge presque-sûrement vers X.*

Il est facile à partir de ce résultat et de la Proposition 12, de montrer que la convergence en probabilité est conservée par les applications continues.

2.2.4 Loi forte des grands nombres

Proposition 15 *Soit (X_k) une suite de variables aléatoires indépendantes, de même loi et intégrables. Alors, $S_n/n \xrightarrow{ps} \mathbb{E}(X_1)$.*

2.3 Convergence en loi

2.3.1 Définition

Définition 9 *Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$. On dit que X_n converge en loi vers X (et on note $X_n \xrightarrow{\mathcal{L}} X$) si pour toute fonction $f : \mathbb{R}^p \to \mathbb{R}$ continue et bornée, $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$).*

Le théorème suivant donne une définition plus pratique (et donc plus usuelle) de la convergence en loi dans le cas $d = 1$.

Théorème 1 *La suite X_n converge en loi vers X si et seulement si $F_n(t) \to F(t)$ lorsque $n \to \infty$ en tout point de continuité de F. On rappelle que $F_n(t) := \mathbb{P}_{X_n}(]-\infty, t])$ et $F(t) = \mathbb{P}_X(]-\infty, t])$.*

Théorème 2 *Soit $h : \mathbb{R}^p \to \mathbb{R}^q$ une fonction continue. Si $X_n \xrightarrow{\mathcal{L}} X$ alors $h(X_n) \xrightarrow{\mathcal{L}} h(X)$.*

Théorème 3 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires de densité f_n et soit X une variable aléatoire de densité f. Si $f_n(t) \to f(t)$ λ-presque-partout alors $X_n \xrightarrow{\mathcal{L}} X$.*

Théorème 4 *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $\mathbb{N}$. La suite X_n converge en loi vers X si et seulement si pour tout $k \in \mathbb{N}$, $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$.*

2.3.2 Théorème de la limite centrale

Théorème 5 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}$ indépendantes, identiquement distribuées et de carré intégrable (i.e. $\mathbb{E}(X_1^2) < \infty$). En posant $\mu = \mathbb{E}(X)$ et $\sigma^2 = \mathrm{Var}(X)$, on a*

$$n^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2).$$

2.3.3 Convergence en loi et convergence en probabilité

Théorème 6 *Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $\mathbb{R}$. Si la suite X_n converge en probabilité vers X alors X_n converge en loi vers X.*

Théorème 7 *Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $\mathbb{R}$. Si $X = a$ presque sûrement et si $X_n \xrightarrow{\mathcal{L}} X$ alors $X_n \xrightarrow{\mathbb{P}} X$.*

2.3.4 Convergence en loi et fonction caractéristique

Rappelons que la fonction caractéristique d'un vecteur aléatoire $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ est donnée par

$$\Phi_X(t) = \mathbb{E}[\exp(i < t, X >)].$$

Théorème 8 *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires de fonctions caractéristique Φ_{X_n} et Φ_X. La suite X_n converge en loi vers X si et seulement si pour tout $t \in \mathbb{R}$, $\Phi_{X_n}(t) \to \Phi_X(t)$.*

2.3.5 Lemme de Slutsky

Proposition 16 (Lemme de Slutsky) *Soient* (X_n) *et* (Y_n) *deux suites de variables aléatoires définies sur l'espace probabilisé* $(\Omega, \mathcal{F}, \mathbb{P})$ *et à valeurs dans* $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$. *Si* $X_n - Y_n \xrightarrow{\mathbb{P}} 0$ *et* $X_n \xrightarrow{\mathcal{L}} X$ *alors* $Y_n \xrightarrow{\mathcal{L}} X$.

Proposition 17 *Pour tout* $n \in \mathbb{N}$, *on considère les variables aléatoires* $X_n : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ *et* $Y_n : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^q, \mathcal{B}(\mathbb{R}^q))$. *Si* $X_n \xrightarrow{\mathcal{L}} X$ *et* $Y_n \xrightarrow{\mathbb{P}} b$ *alors* $(X_n, Y_n) \xrightarrow{\mathcal{L}} (X, b)$.

2.4 Convergence dans L^p

Définition 10 *Soient* (X_n) *et* X *des variables aléaoires admetttant un moment d'ordre* p. *On dit que* X_n *converge dans* $L^p(\Omega, \mathcal{F}, \mathbb{P})$ *vers* X *si*

$$\lim_{n\to\infty} \mathbb{E}\left(|X_n - X|^p\right) = 0.$$

Cette convergence est notée $X_n \xrightarrow{L^p} X$.

Proposition 18 *Si* X_n *converge en moyenne vers* X *alors* $\mathbb{E}(X_n) \to \mathbb{E}(X)$ *et* $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$.

Théorème 9 *S'il existe* $p \in \mathbb{N} \setminus \{0\}$ *tel que* $X_n \xrightarrow{L^p} X$ *alors* $X_n \xrightarrow{\mathbb{P}} X$.

Diagramme des convergences stochastiques

$$\begin{array}{ccccc} \xrightarrow{ps} & \Rightarrow & \xrightarrow{\mathbb{P}} & \Rightarrow & \xrightarrow{\mathcal{L}} \\ & & \Uparrow & & \\ & & \xrightarrow{L^p} & & \end{array}$$