Probabilités 4
Double Licence Mathématiques - Economie
3ème année

Table des matières

I Convergence d'une suite de variables aléatoires 3

1 Convergence en probabilité 4
1.1 Définition et propriétés 4
1.2 Loi faible des grands nombres 5

2 Convergence presque-sûre 6
2.1 Définitions 6
2.2 Lemme de Borel-Cantelli 6
2.3 Convergence en probabilité et convergence presque-sûre 7
2.4 Loi forte des grands nombres 7

3 Convergence en loi 8
3.1 Définition et principaux résultats 8
3.2 Théorème de la limite centrale 9
3.3 Convergence en loi et convergence en probabilité 9
3.4 Convergence en loi et fonction caractéristique 9

4 Convergence dans L^p 10

II Espérance conditionnelle 11

5 Définition et propriétés 12
5.1 Cas général 12
5.2 Cas des variables aléatoires de carré intégrable 13

6 Espérance conditionnelle par rapport à une variable aléatoire 14
6.1 Conditionnement par rapport à une variable aléatoire discrète . . 14
6.2 Conditionnement par rapport à une variable aléatoire à densité . 14

Première partie

Convergence d'une suite de variables aléatoires

Chapitre 1

Convergence en probabilité

1.1 Définition et propriétés

Définition 1 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire et à valeurs dans $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. On dit que X_n converge en probabilité vers X si*

$$\forall\varepsilon>0,\ \lim_{n\to\infty}\mathbb{P}(|X_n-X|\geq\varepsilon)=0\Leftrightarrow X_n\xrightarrow{\mathbb{P}}X.$$

La définition de la convergence en probabilité de variables aléatoires réelles s'étend bien sûr aux suites de vecteurs aléatoires $(X_n)_{n\in\mathbb{N}}=(X_{n,1},\ldots,X_{n,p})^\top_{n\in\mathbb{N}}$.

Définition 2 *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.*

Nous allons à présent montrer que la convergence en probabilité est une notion topologique c'est-à-dire qu'elle est associée à une distance sur l'ensemble $L^0(\Omega,\mathcal{F},\mathbb{P})$ des classes d'équivalence pour l'égalité presque-sûre des variables aléatoires. On introduit tout d'abord la distance $d(\cdot,\cdot)$ sur $L^0(\Omega,\mathcal{F},\mathbb{P})$ donnée par

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

Proposition 1 *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 2 *Si $(X_n)_{n\in\mathbb{N}}$ converge en probabilité alors sa limite est unique presque-sûrement.*

Proposition 3 *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 $f(X_n)\xrightarrow{\mathbb{P}}f(X)$.*

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

Chapitre 2

Convergence presque-sûre

2.1 Définitions

Définition 3 *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 *si*

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

Définition 4 *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 Lemme de Borel-Cantelli

Proposition 5 (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,*

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

L'utilisation principale de ce résultat est donnée par le corollaire suivant.

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 6 *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 $f(X_n) \xrightarrow{ps} f(X)$.*

2.3 Convergence en probabilité et convergence presque-sûre

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

Proposition 7 *Si $X_n \xrightarrow{ps} X$ alors $X_n \xrightarrow{\mathbb{P}} X$.*

Proposition 8 *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.*

On déduit de ce résultat et de la Proposition 6 que la convergence en probabilité est conservée par les applications continues.

2.4 Loi forte des grands nombres

Proposition 9 *Soit (X_k) une suite de variables aléatoires indépendantes, de même loi et intégrables. Alors, $S_n/n \xrightarrow{ps} \mathbb{E}(X_1)$.*

Chapitre 3

Convergence en loi

3.1 Définition et principaux résultats

Définition 5 *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* $\mathbb{P}_{X_n}$ *converge étroitement vers* $\mathbb{P}_X$ *(autrement dit, si pour toute fonction* $f : \mathbb{R}^p \to \mathbb{R}$ *continue et bornée,* $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$*).*

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* $F(t) = \mathbb{P}_X(]-\infty, t])$.

Le théorème suivant montre que la convergence en loi est conservée par les applications continues.

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

Le résultat suivant est dédié aux variables aléatoires réelles admettant une densité par rapport à la mesure de Lebesgue.

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* $X_n \xrightarrow{\mathcal{L}} X$.

Le résultat suivant donne une condition nécessaire et suffisante pour la convergence en loi de variables aléatoires discrètes. Pour simplifier la rédaction, on supposera que la variable aléatoire prend ses valeurs dans $\mathbb{N}$.

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* $k \in \mathbb{N}$, $\mathbb{P}(X_n = k) \to \mathbb{P}(X = k)$.

3.2 Théorème de la limite centrale

Le théorème ci-dessous (que l'on admettra) est un résultat très utile en statistique.

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*

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

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.*

La réciproque du théorème 6 est fausse en général car la convergence en loi ne requiert pas que les variables aléatoires soient définies sur le même espace probabilisé au contraire de la convergence en probabilité. Elle est cependant vraie dans le cas particulier ci-dessous.

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 $X_n \xrightarrow{\mathbb{P}} X$.*

3.4 Convergence en loi et fonction caractéristique

Le théorème suivant permet de faire le lien entre fonction caractéristique et convergence en loi.

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 $t \in \mathbb{R}$, $\Phi_{X_n}(t) \to \Phi_X(t)$.*

Chapitre 4

Convergence dans L^p

On rappelle qu'une variable aléatoire $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ admet un moment d'ordre $p \in \mathbb{N} \setminus \{0\}$ si $\mathbb{E}(|X|^p) < \infty$. On note $L^p(\Omega, \mathcal{F}, \mathbb{P})$ l'ensemble des classes (pour l'équivalence presque-sûre) des variables aléatoires admettant un moment d'ordre p.

Définition 6 *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 10 *Si X_n converge en moyenne vers X alors $\mathbb{E}(X_n) \to \mathbb{E}(X)$ et $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$.*

Voici maintenant un résultat faisant le lien entre convergence en probabilité et convergence dans L^p.

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

On conclut cette partie par le diagramme des convergences stochastiques rappelant les différentes implications entre les quatre types de convergence.

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

Deuxième partie

Espérance conditionnelle

Chapitre 5

Définition et propriétés

5.1 Cas général

Définition 7 *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 8 *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 11 *L'espérance conditionnelle de* X *sachant* $\mathcal{A}$ *existe et est unique (modulo l'égalité presque-sûre).*

Proposition 12 *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*

$$\mathbb{E}(X|\mathcal{A}) = \sum_{n\in\mathbb{N}} \left(\frac{1}{\mathbb{P}(A_n)} \int_{A_n} X d\mathbb{P} \right) \mathbb{I}_{A_n}.$$

5.2 Cas des variables aléatoires de carré intégrable

Le résultat suivant assure que dans le cas particulier où la variable X admet un moment d'ordre 2, l'espérance conditionnelle coïncide avec la projection orthogonale sur $L^2(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition 13 *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,*

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

Chapitre 6

Espérance conditionnelle par rapport à une variable aléatoire

Définition 9 *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 $\mathbb{E}(X|\sigma(Y))$.*

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 $Z = h(Y)$.*

6.1 Conditionnement par rapport à une variable aléatoire discrète

$$\mathbb{E}(X|Y) = \sum_{n \geq 1} \left(\frac{1}{\mathbb{P}(Y^{-1}(\{x_n\}))} \int_{Y^{-1}(\{x_n\})} X d\mathbb{P} \right) \mathbb{I}_{Y^{-1}(\{x_n\})}.$$

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

$$\mathbb{E}(X|Y) = \sum_{n \geq 1} \left(\frac{1}{\mathbb{P}(Y = x_n)} \int_{\{Y = x_n\}} X d\mathbb{P} \right) \mathbb{I}_{\{Y = x_n\}}.$$

6.2 Conditionnement par rapport à une variable aléatoire à densité

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.

Définition 10 *La densité conditionnelle de X sachant $Y = y$ est la fonction mesurable positive définie par :*

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

Ce dernier résultat permet de calculer facilement des espérances conditionnelles.

Proposition 14 *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*

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