Appendice. Quelques prérequis de probabilités niveau L3-M1

Courte bibliographie

Il existe une multitude d'ouvrages de probabilités, au niveau L3-M1. Nous vous recommendons par exemple les ouvrages suivants :

- **Barbe-Ledoux, Probabilité L3-M1** *[bases de théorie de la mesure, convergence de v.a., vecteurs gaussiens, espérance conditionnelle]*
- **Jacod-Protter, L'essentiel en théorie des probabilités** *[convergence de v.a., vecteurs gaussiens, espérance conditionnelle, martingales à temps discret]*

Pour la théorie des martingales à temps continu, nous vous recommandons :

- **Le Gall, Mouvement brownien, martingales et calcul stochastique** *[chapitre 3]*

Comme d'habitude, dans tout l'appendice, on se place sur un espace de probabilités $(\Omega, \mathcal{F}, P)$.

1. Espérance conditionnelle

Dans tous les énoncés, l'unicité s'entend "à un ensemble de mesure nulle près". Toutes les égalités ou inégalités faisant intervenir des espérances conditionnelles sont à comprendre au sens presque sûr.

1.1. Définition pour des variables dans $\mathbf{L^1(\Omega, \mathcal{F}, P)}$

Théorème-Définition 1

*Soit $\mathcal{G}$ une sous tribu de $\mathcal{F}$ et $X \in L^1(\Omega, \mathcal{F}, P)$. Il existe une unique variable aléatoire dans $L^1(\Omega, \mathcal{G}, P)$ (donc $\mathcal{G}$-mesurable), notée $E_{\mathcal{G}}(X)$ (ou $E(X|\mathcal{G})$ et appelée **espérance conditionnelle** de X sachant $\mathcal{G}$, telle que*

$$\forall B \in \mathcal{G}, \quad E(X\mathbf{1}_B) = E(E_{\mathcal{G}}(X)\mathbf{1}_B).$$

*La relation ci-dessus est appelée **propriété caractéristique de l'espérance conditionnelle**.*
On a plus généralement que pour toute v.a. Z $\mathcal{G}$-mesurable bornée,

$$E(XZ) = E(E_{\mathcal{G}}(X)Z).$$

Si $X \geq 0$, on a aussi $E_{\mathcal{G}}(X) \geq 0$.

Dans le cas où la sous-tribu $\mathcal{G}$ est la tribu $\sigma(Y)$ engendrée par une v.a. Y, on notera l'espérance conditionnelle correspondante $E_{\sigma(Y)}(X)$ ou $E_Y(X)$ (ou encore $E(X|Y)$).

1.2. Pour des variables positives

Théorème 2

Soit X une v.a. à valeurs dans $[0, \infty]$.
La formule $E_{\mathcal{G}}(X) := \lim\limits_{n\to\infty} E_{\mathcal{G}}(\inf(X, n))$ (où la limite est croissante) définit une v.a. à valeurs dans $[0, \infty]$ qui est caractérisée par la propriété suivante :

$$\forall Z \text{ } \mathcal{G}\text{-mesurable positive, } E(XZ) = E(E_{\mathcal{G}}(X)Z).$$

1.3. Propriétés de l'espérance conditionnelle

Proposition 3

On suppose que X et X' sont positives ou dans $L^1(\Omega, \mathcal{F}, P)$. Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.

1. *(linéarité) Pour tout $a, b \in \mathbb{R}$, $E_{\mathcal{G}}(aX + bX') = aE_{\mathcal{G}}(X) + bE_{\mathcal{G}}(X')$.*
2. *Si X est $\mathcal{G}$-mesurable, $E_{\mathcal{G}}(X) = X$.*
3. *(convergence monotone) Si $(X_n)_{n\in\mathbb{N}}$ est une suite croissante de v.a. positives et X la limite croissante des X_n, alors*
$$E_{\mathcal{G}}(X) = \lim_{n\to\infty} E_{\mathcal{G}}(X_n),$$
où la limite est croissante.
4. *(Fatou) Si $(X_n)_{n\in\mathbb{N}}$ est une suite de v.a. positives, alors*
$$E_{\mathcal{G}}(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} E_{\mathcal{G}}(X_n).$$
5. *(convergence dominée) Soit $(X_n)_{n\in\mathbb{N}}$ une suite de v.a. intégrables qui converge p.s. vers X. Suppposons qu'il existe une v.a. Z telle que $\forall n$, $|X_n| \leq Z$ p.s. et $E(Z) < \infty$. Alors*
$$E_{\mathcal{G}}(X) = \lim_{n\to\infty} E_{\mathcal{G}}(X_n), \quad \text{p.s. et dans } L^1.$$

6. *Si $X \in L^1(\Omega, \mathcal{F}, P)$, $E(E_{\mathcal{G}}(X)) = E(X)$.*
7. *Si $X \in L^1(\Omega, \mathcal{F}, P)$, $|E_{\mathcal{G}}(X)| \leq E_{\mathcal{G}}(|X|)$ et donc $E(|E_{\mathcal{G}}(X)|) \leq E(|X|)$.*
8. *(Jensen) Si f est une fonction convexe positive ou telle que $f(X) \in L^1$, alors*
$$E_{\mathcal{G}}(f(X)) \geq f(E_{\mathcal{G}}(X)).$$
9. *Si Y est $\mathcal{G}$-mesurable, avec X et Y positives ou bien X et XY dans L^1, alors*
$$E_{\mathcal{G}}(XY) = YE_{\mathcal{G}}(X).$$
10. *Si $\mathcal{G}_1$ et $\mathcal{G}_2$ sont deux sous-tribus de $\mathcal{F}$ telles que $\mathcal{G}_1 \subset \mathcal{G}_2$, $E_{\mathcal{G}_1}(E_{\mathcal{G}_2}(X)) = E_{\mathcal{G}_1}(X) = E_{\mathcal{G}_2}(E_{\mathcal{G}_1}(X))$.*
11. *Soient X est une variable aléatoire $\mathcal{G}$-mesurable et Y une variable aléatoire indépendante de $\mathcal{G}$. Alors pour toute fonction Φ mesurable, positive ou bornée,*
$$p.s. \quad E_{\mathcal{G}}(\Phi(X,Y)) = \varphi(X),$$
avec pour tout x, $\varphi(x) = E(\Phi(x,Y))$.

NB : les propriétés 6., 8., 9. et 11. sont particulièrement utiles dans les calculs.

1.4. Espérance conditionnelle pour des variables de carré intégrable

On appelle $L^2(\Omega, \mathcal{G}, P)$ le sous-espace fermé des éléments de $L^2(\Omega, \mathcal{F}, P)$ dont au moins un représentant est $\mathcal{G}$-mesurable.

Théorème 4

Si $X \in L^2(\Omega, \mathcal{F}, P)$, alors $E_{\mathcal{G}}(X)$ est la projection orthogonale de X sur $L^2(\Omega, \mathcal{G}, P)$, pour le produit scalaire $(X,Y) \mapsto \mathrm{E}(XY)$.

---

2. Vecteurs gaussiens

Soit C une matrice de taille $d \times d$ symétrique positive, à coefficients dans $\mathbb{R}$. Un **vecteur gaussien centré de matrice de covariance C** est un vecteur aléatoire X à valeurs dans $\mathbb{R}^d$, dans L^2, dont la fonction caractéristique est donnée par

$$\varphi_X(t_1, \ldots, t_d) = E(e^{i\langle t, X\rangle}) = \exp\left(-\frac{1}{2}\sum_{j,k=1}^{d} C_{jk}t_jt_k\right), \quad \forall t = (t_1, \ldots, t_d) \in \mathbb{R}^d.$$

On dit alors que X suit une loi $\mathcal{N}(0,C)$. Cette notation est justifiée par le fait que son espérance est nulle et sa matrice de covariance est C.

Pour toute matrice réelle symétrique, positive, il existe un vecteur gaussien centré de matrice de covariance C, obtenu comme AY, où $Y = (Y^1, \ldots, Y^d)$ un vecteur aléatoire dont les coordonnées sont des variables aléatoires réelles indépendantes, toutes de loi $\mathcal{N}(0,1)$ et $A = \sqrt{C}$.

Un vecteur aléatoire X à valeurs dans $\mathbb{R}^d$ est un vecteur gaussien centré si et seulement si toute combinaison linéaire de ses coordonnées est une variable aléatoire (réelle) gaussienne centrée.

Soit X un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, de loi $\mathcal{N}(0, C)$. Si C n'est pas inversible, la loi de X n'a pas de densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Si C est inversible, la densité de la loi de X est donnée par, $\forall x \in \mathbb{R}^d$

$$f_X(x) = \frac{1}{(2\pi)^{d/2}\sqrt{det(C)}} exp\left(-\frac{1}{2}\langle x, C^{-1}x\rangle\right).$$

Proposition 5

Soit $(X_1, \ldots, X_m, Y_1, \ldots, Y_n)$ un vecteur gaussien centré. Alors les vecteurs $(X_1, \ldots, X_m)$ et $(Y_1, \ldots, Y_n)$ sont indépendants si et seulement si, pour tout couple (i, j) avec $i \in \{1, \ldots, m\}$ et $j \in \{1, \ldots, n\}$, $cov(X_i, Y_j) = 0$.

Pour les vecteurs gaussiens, les calculs d'espérance conditionnelle se ramènent à une projection orthogonale, pour le produit scalaire $(X, Y) \mapsto \mathrm{E}(XY)$

Proposition 6

Soit $(Y_1, \ldots, Y_n, X)$ un vecteur gaussien centré. Alors l'espérance conditionnelle de X sachant $(Y_1, \ldots, Y_n)$ est donnée par

$$E[X|Y_1, \ldots, Y_n] = \hat{X},$$

où $\hat{X}$ est la projection orthogonale de X sur l'espace vectoriel $\{\sum_{k=1}^{n} \alpha_k Y_k; \alpha_k \in \mathbb{R}\}$.

---

3. Les différentes notions de convergence pour des v.a. réelles

On dit que la suite de v.a. $(X_n)_{n\geq 1}$ **converge presque sûrement** (p.s.) vers la v.a. X si

$$P\left(\lim_{n\to+\infty} X_n = X\right) = 1.$$

Autrement dit, il existe un sous-ensemble Ω' tel que $P(\Omega') = 1$ et $\forall \omega \in \Omega'$, $\lim\limits_{n\to+\infty} X_n(\omega)$ existe et vaut $X(\omega)$. On note $X_n \xrightarrow{p.s.} X$.

Proposition 7 *(un critère pour la convergence p.s.)*

1. *Si $\forall \epsilon > 0$, $\sum_{n\geq 1} P(|X_n - X| > \epsilon) < \infty$, alors $X_n \xrightarrow{p.s.} X$.*
2. *On suppose que les $(X_n)_{n\geq 1}$ sont indépendants. Alors $X_n \xrightarrow{p.s.} 0$ ssi $\forall \epsilon > 0$, $\sum_{n\geq 1} P(|X_n| > \epsilon) < \infty$.*

Proposition 8 *(une autre caractérisation de la convergence p.s.)*

> *La suite $(X_n)_{n\geq 1}$ converge p.s. vers X ssi*
>
> $$\forall \epsilon > 0, \lim_{n\to\infty} P\left(\sup_{k\geq n} |X_k - X| \geq \epsilon\right) = 0.$$

On dit que la suite $(X_n)_{n\geq 1}$ **converge en probabilité** vers la v.a. X si

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

On note $X_n \xrightarrow{P} X$.

Proposition 9 *(lien entre convergence p.s. et convergence en probabilité)*

> *Si $X_n \xrightarrow{p.s.} X$ alors $X_n \xrightarrow{P} X$.*
>
> *La réciproque est fausse en général.*
>
> *Si $X_n \xrightarrow{P} X$ alors il existe une sous-suite extraite $(X_{n_k})_{k\geq 1}$ qui converge p.s. vers X.*

La convergence en probabilité est stable par les opérations usuelles :

Proposition 10

> *On suppose $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$ et φ continue de $\mathbb{R}^2$ dans $\mathbb{R}$ alors $\varphi(X_n, Y_n) \xrightarrow{P} \varphi(X, Y)$.*
>
> *En particulier, si $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, alors pour touts réels α et β, $\alpha X_n + \beta Y_n \xrightarrow{P} \alpha X + \beta Y$, ou encore $X_n Y_n \xrightarrow{P} XY$, etc.*

Soit $L^0(\Omega, \mathcal{F}, P)$ l'ensemble des v.a. réelles sur $(\Omega, \mathcal{F}, P)$, quotienté par la relation d'équivalence $X \sim Y$ ssi $X = Y$ p.s. Pour tous $X, Y \in L^0(\Omega, \mathcal{F}, P)$, on définit $d(X, Y) = E(min(|X - Y|, 1))$.

Proposition 11 *(métrisabilité de la convergence en probabilité)*

> *d est une distance sur $L^0(\Omega, \mathcal{F}, P)$ et $(X_n)_{n\geq 1}$ converge en probabilité vers X ssi $d(X_n, X)$ converge vers 0. On dit que d métrise la convergence en probabilité.*

On peut même montrer que $L^0(\Omega, \mathcal{F}, P)$ muni de la distance d est un espace métrique complet :

Proposition 12 *(complétude de $L^0(\Omega, \mathcal{F}, P)$)*

> *Supposons que $(X_n)_{n\geq 1}$ vérifie le critère de Cauchy pour la distance d, i.e. $\forall \epsilon > 0, \exists n_0, \forall n \geq n_0$, $d(X_n, X_{n_0}) \leq \epsilon$. Alors $(X_n)_{n\geq 1}$ converge en probabilité.*

La suite de v.a. $(X_n)_{n\geq 1}$ **converge dans** L^p vers la v.a. X si $\lim_{n\to+\infty} \|X_n - X\|_p = 0$, autrement dit

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

On note $X_n \xrightarrow{L^p} X$.

Les plus usitées sont les convergences L^1 ou L^2.

Il est facile de vérifier par Jensen que s'il y a convergence L^p, il y a convergence $L^{p'}$ pour tout $0 < p' < p$.

Proposition 13 *(lien avec la convergence en probabilité)*

> *Si $X_n \xrightarrow{L^p} X$ alors $X_n \xrightarrow{P} X$.*
>
> *La réciproque est fausse en général.*

Pour établir une réciproque partielle, on rappelle la notion d'uniforme intégrabilité :

Une famille $(X_i)_{i\in I}$ de v.a. réelles, intégrables, est dite équiintégrable ou uniformément intégrable si

$$\lim_{c\to\infty} \sup_{i\in I} \int_{\{|X_i|>c\}} |X_i| dP = 0.$$

On rappelle quelques critères d'uniforme intégrabilité :
- une famille finie de v.a. intégrables est uniformément intégrable,
- si il existe une v.a. Y intégrable telle que, p.s. $\forall i \in I, |X_i| \leq Y$, alors la famille $(X_i)_{i\in I}$ est uniformément intégrable,
- la famille $(X_i)_{i\in I}$ est uniformément intégrable ssi ($\sup_{i\in I} E(|X_i|) < \infty$ et $\forall \epsilon > 0, \exists \eta > 0, \forall A$ telle que $P(A) \leq \eta, \forall i \in I, \int_A |X_i| dP \leq \epsilon$.

Proposition 14 *(lien entre les convergences L^1 et en probabilité)*

> *Soit $(X_n)_{n\geq 1}$ une suite de v.a. intégrables. Les deux propositions suivantes sont équivalentes :*
>
> 1. *$X_n \xrightarrow{P} X$ et la famille $(X_n)_{n\geq 1}$ est équiintégrable.*
> 2. *X est intégrable et $\lim_{n\to\infty} \|X_n - X\|_1 = 0$.*

Remarque : La convergence L^p n'implique pas la convergence p.s., ni l'inverse.

On ne suppose plus nécessairement que les v.a. $(X_n)_{n\geq 1}$ sont définies sur le même espace de probabilités. La convergence en loi correspond à la convergence étroite des lois des v.a. $(X_n)_{n\geq 1}$.

Théorème 15 *(définition et caractérisation de la convergence en loi)*

> *Soient $(X_n)_{n\geq 1}$ et X des v.a. réelles. On dit que $(X_n)_{n\geq 1}$ converge en loi vers X si l'une de ces trois conditions équivalentes est vérifiée :*
>
> 1. *$\forall \varphi : \mathbb{R} \to \mathbb{R}$ continue bornée,*
> $$\lim_{n\to\infty} E(\varphi(X_n)) = E(\varphi(X)).$$

2. *Soient F_{X_n} et F_X respectivement les fonctions de répartition de X_n et de X. En tout point de continuité t de F_X, on a*
$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t)$$
3. *Il existe un espace de probabilité $(\Omega', \mathcal{F}', P')$ sur lequel sont définies des v.a. $(X'_n)_{n\geq 1}$ et X' de même loi respectivement que $(X_n)_{n\geq 1}$ et X telles que $X'_n \xrightarrow{p.s.} X'$.*

On note alors $X_n \xrightarrow{\mathcal{L}} X$.
Attention, pas d'opérations usuelles sur la convergence en loi.

Proposition 16

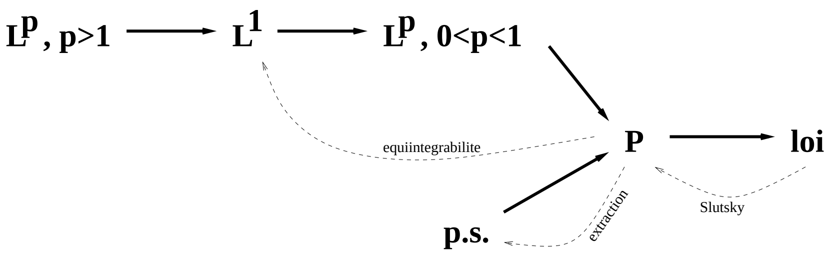
Si $X_n \xrightarrow{p.s.} X$ alors $X_n \xrightarrow{\mathcal{L}} X$.
Si $X_n \xrightarrow{P} X$ alors $X_n \xrightarrow{\mathcal{L}} X$.

Les réciproques sont fausses

Théorème 17 *(Lévy)*

Soient $(X_n)_{n\geq 1}$ et X des v.a. réelles.
1. *Si $X_n \xrightarrow{\mathcal{L}} X$, alors $\lim_{n\to+\infty} E\left[e^{itX_n}\right] = E\left[e^{itX}\right]$ pour tout $t \in \mathbb{R}$.*
2. *Si $\lim_{n\to+\infty} E\left[e^{itX_n}\right] = \varphi(t)$ pour tout $t \in \mathbb{R}$ et φ est la fonction caractéristique d'une v.a. X, alors $X_n \xrightarrow{\mathcal{L}} X$.*

En résumé :



Sur le schéma ci-dessous, les flèches pleines indiquent les liens qui ont toujours lieu et les flèches en pointillés indiquent les réciproquent partielles.

4. Martingales à temps discret

Dans cette partie, on se place dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ muni d'une suite croissante $(\mathcal{F}_n)_{n\in\mathbb{N}}$ de sous-tribus de $\mathcal{F}$.
Définition.

Soit $(X_n)_{n\geq 0}$ un processus $(\mathcal{F}_n)$-adapté tel que, pour tout $n \geq 0$, $X_n \in \mathbb{L}^1$.

- $(X_n)_{n\geq 0}$ est une martingale ssi pour tout $n \geq m$, $X_m = E(X_n|\mathcal{F}_m)$.
- $(X_n)_{n\geq 0}$ est une sous-martingale ssi pour tout $n \geq m$, $X_m \leq E(X_n|\mathcal{F}_m)$.
- $(X_n)_{n\geq 0}$ est une sur-martingale ssi pour tout $n \geq m$, $X_m \geq E(X_n|\mathcal{F}_m)$.

Proposition 18 *[Inégalité de Doob]*

Soit $(M_n)_{n\geq 0}$ une sous-martingale positive ou une martingale.

$$\mathrm{E}\left[\sup_{k\leq n} M_k^2\right] \leq 4\mathrm{E}[M_n^2].$$

L'une des nombreuses versions du théorèmes d'arrêt de Doob peut s'énoncer de la façon suivante :

Théorème 19 *[Théorème d'arrêt de Doob]*

Si $(M_n)_{n\geq 0}$ est une martingale et T un temps d'arrêt, alors le processus M^T défini par $M_n^T = M_{T\wedge n}$ pour tout $n \in \mathbb{N}$ est encore une martingale.

Théorème 20 *[Théorèmes de convergence pour les martingales]*

Soit $(M_n)_{n\geq 0}$ une martingale bornée dans $\mathbb{L}^1$. Alors $(M_n)_{n\geq 0}$ converge p.s. vers une variable que l'on note M_∞. Les propositions suivantes sont équivalentes :

1. *$(M_n)_{n\geq 0}$ converge dans $\mathbb{L}^1$ vers M_∞.*
2. *$(M_n)_{n\geq 0}$ est uniformément intégrable.*
3. *Pour tout $n \geq 0$, $M_n = E(M_\infty|\mathcal{F}_n)$.*
4. *$\exists M \in \mathbb{L}^1$, tel que $M_n = E(M_\infty|\mathcal{F}_n)$. On a alors $M_\infty = E(M|\mathcal{F}_\infty)$.*

5. Martingales à temps continu

Dans cette partie, on se place dans un espace de probabilité $(\Omega, \mathcal{F}, P)$ muni d'une suite croissante $(\mathcal{F}_t)_{t\geq 0}$ de sous-tribus de $\mathcal{F}$. On suppose la filtration $(\mathcal{F}_t)_{t\geq 0}$ continue à droite et complète. Des résultats analogues à ceux à temps discret sont vrais pour pour des martingales à trajectoires continues.
Définition.

Soit $(X_t)_{t\geq 0}$ un processus $(\mathcal{F}_t)$-adapté tel que, pour tout $t \geq 0$, $X_t \in \mathbb{L}^1$.

- $(X_t)_{t\geq 0}$ est une martingale ssi pour tout $t \geq s$, $X_s = E(X_t|\mathcal{F}_s)$.
- $(X_t)_{t\geq 0}$ est une sous-martingale ssi pour tout $t \geq s$, $X_s \leq E(X_t|\mathcal{F}_s)$.
- $(X_t)_{t\geq 0}$ est une sur-martingale ssi pour tout $t \geq s$, $X_s \geq E(X_t|\mathcal{F}_s)$.

Proposition 21 *[Inégalité de Doob]*

Soit $(M_t)_{t\geq 0}$ une martingale à trajectoires continues. Pour tout $T > 0$,

$$\mathrm{E}\left[\sup_{t\leq T} M_t^2\right] \leq 4\mathrm{E}[M_T^2].$$

L'une des nombreuses versions du théorèmes d'arrêt de Doob peut s'énoncer de la façon suivante :

Théorème 22 *[Théorème d'arrêt de Doob]*

Si $(M_t)_{t\geq 0}$ est une martingale à trajectoires continues et T un temps d'arrêt, alors le processus M^T défini par $M_t^T = M_{T\wedge t}$ pour tout $t \geq 0$ est encore une martingale.

Théorème 23 *[Théorèmes de convergence pour les martingales]*

Soit $(M_t)_{t\geq 0}$ une martingale à trajectoires continues, bornée dans $\mathbb{L}^1$. Alors $(M_t)_{t\geq 0}$ converge p.s. vers une variable que l'on note M_∞. Les propositions suivantes sont équivalentes :

1. *$(M_t)_{t\geq 0}$ converge dans $\mathbb{L}^1$ vers M_∞.*
2. *$(M_t)_{t\geq 0}$ est uniformément intégrable.*
3. *Pour tout $t \geq 0$, $M_t = E(M_\infty|\mathcal{F}_t)$.*
4. *$\exists M \in \mathbb{L}^1$, tel que $M_t = E(M_\infty|\mathcal{F}_t)$. On a alors $M_\infty = E(M|\mathcal{F}_\infty)$.*