CHAPITRE 0: QUELQUES RAPPELS DE THÉORIE DES PROBABILITÉS

MR2, module B1: Cours de calcul stochastique
Fabrice Baudoin

Sans être exhaustif, nous rappelons ici quelques éléments de la théorie des probabilités qui reviendront tout au long de ce cours. Des notions telles que:

- La notion d'espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$;
- La notion de variable aléatoire;
- La notion de loi d'une variable aléatoire;
- La notion de fonction caractéristique;
- Le théorème central limite et la loi des grands nombres;
- La notion d'indépendance de tribus ou de variables aléatoires;
- La notion de loi gaussienne et de vecteurs gaussiens;

sont supposées acquises. Dans le cas contraire, nous renvoyons à tout ouvrage élémentaire dont notamment:

"*Probability*" par A.N. Shiryaev, Springer, Graduate Texts in Mathematics, 95.

Table des matières

1. Lemme de Borel-Cantelli

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité. Soit $(A_n)_{n\in\mathbb{N}}$, une suite d'événements de $\mathcal{F}$. On note

$$\limsup A_n = \bigcap_m \bigcup_{n\geq m} A_n = \{\omega\in\Omega,\forall m,\exists n(\omega)\geq m,\omega\in A_{n(\omega)}\}.$$

Proposition 1. *(Lemme de Borel-Cantelli) Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de $\mathcal{F}$ telle que*

$$\sum_n \mathbb{P}(A_n) < +\infty.$$

Alors

$$\mathbb{P}(\limsup A_n) = 0.$$

2. Quelques inégalités

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité.

Proposition 2. *(Inégalité de Markov) Soient X une variable aléatoire mesurable et $g:\mathbb{R}\rightarrow(0,+\infty)$ une fonction borélienne, alors pour tout $x\in\mathbb{R}$,*

$$\mathbb{P}(X\geq x)\leq\frac{\mathbb{E}(g(X))}{g(x)}.$$

Proposition 3. *(Inégalité de Jensen) Soient $\phi:\mathbb{R}\rightarrow\mathbb{R}$ une fonction convexe et X une variable aléatoire mesurables. Supposons que,*

$$\mathbb{E}(\mid X\mid)<+\infty,\ \mathbb{E}(\mid\phi(X)\mid)<+\infty.$$

Alors,

$$\mathbb{E}(\phi(X))\geq\phi(\mathbb{E}(X)).$$

Proposition 4. *(Inégalité de Tchebychev) Soit X une variable aléatoire mesurable dans L^2, alors pour tout $c>0$,*

$$\mathbb{P}(\mid X-\mathbb{E}(X)\mid\geq c)\leq\frac{\mathbf{Var}(X)}{c^2}.$$

Proposition 5. *(Inégalités de Hölder et Minkowski) Soient $p>1$ et q tels que $\frac{1}{p}+\frac{1}{q}=1$. On a pour toutes variables aléatoires mesurables X et Y,*

$$\mid\mathbb{E}(XY)\mid\mathbb{E}(\mid X\mid^p)^{\frac{1}{p}}\mathbb{E}(\mid Y\mid^q)^{\frac{1}{q}},$$

et

$$\mathbb{E}(\mid X+Y\mid^p)^{\frac{1}{p}}\leq\mathbb{E}(\mid X\mid^p)^{\frac{1}{p}}+\mathbb{E}(\mid Y\mid^p)^{\frac{1}{p}}.$$

3. Convergence d'une suite de variables aléatoires

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles mesurables.

Definition 6. *On dit que $X_n\rightarrow_{n\rightarrow+\infty}X$ presque sûrement si*

$$\mathbb{P}(X_n\rightarrow_{n\rightarrow+\infty}X)=1.$$

On dit que $X_n\rightarrow_{n\rightarrow+\infty}X$ en probabilité si pour tout $\varepsilon>0$,

$$\mathbb{P}(\mid X_n-X\mid>\varepsilon)\rightarrow_{n\rightarrow+\infty}0.$$

On dit $X_n\rightarrow_{n\rightarrow+\infty}X$ dans L^p, $p\geq1$, si

$$\mathbb{E}(\mid X_n-X\mid^p)^{\frac{1}{p}}\rightarrow_{n\rightarrow+\infty}0.$$

Entre ces différents modes de convergence, on a les relations suivantes:

- La convergence presque sûre implique la convergence en probabilité;
- La convergence dans L^p implique la convergence en probabilité.

D'autre part,

Proposition 7. *(Théorème de convergence dominée) Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles mesurables telle que $X_n \to_{n\to+\infty} X$ en probabilité. Si il existe une constante $K>0$ telle que*

$$\sup_n \mid X_n \mid \leq K,$$

alors $X_n \to_{n\to+\infty} X$ presque sûrement.

Definition 8. *Soit $(X_i)_{i\in\mathcal{I}}$ une famille de variables aléatoires. On dit que $(X_i)_{i\in\mathcal{I}}$ est uniformément intégrable si pour tout $\varepsilon>0$, on peut trouver un $K\geq 0$ tel que*

$$\forall i \in \mathcal{I}, \quad \mathbb{E}(\mid X_i \mid 1_{\mid X_i \mid > K}) < \varepsilon.$$

On a alors quelques propriétés:

- Une famille finie de L^1 est toujours uniformément intégrable;
- Une famille uniformément intégrable est bornée dans L^1;
- Une famille bornée dans L^p, $p>1$, est uniformément intégrable.

Proposition 9. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires appartenant à L^1. Soit $X\in L^1$. Alors $X_n \to_{n\to+\infty} X$ dans L^1, si et seulement si:*

1. *$X_n \to_{n\to+\infty} X$ en probabilité;*
2. *La suite $(X_n)_{n\in\mathbb{N}}$ est uniformément intégrable.*

La notion de convergence en loi est la plus faible mais aussi la plus intéressante d'un point de vue probabiliste.
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires mesurables à valeurs dans un espace polonais[1].

Definition 10. *On dit que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire X si pour toute fonction $f: E\to\mathbb{R}$, continue, bornée, on a:*

$$\mathbb{E}(f(X_n)) \to_{n\to+\infty} \mathbb{E}(f(X_n)).$$

Cela revient à dire que la suite des lois de X_n converge étroitement vers la loi de X.

On remarquera que pour la convergence en loi, les variables aléatoires X_n n'ont donc pas nécessairement besoin d'être définies sur le même espace de probabilité.

Théorème 11. *(Théorème de Prokhorov) Soit $\mathcal{P}$ une famille de probabilités définies sur $(E,\mathbb{E})$. Alors $\mathcal{P}$ est un ensemble relativement compact pour la topologie de la convergence étroite, si et seulement si la famille $\mathcal{P}$ est tendue, i.e. pour tout $\varepsilon\in(0,1)$, on peut trouver un ensemble compact $K_\varepsilon\subset E$ tel que pour tout $\mathbb{P}\in\mathcal{P}$,*

$$\mathbb{P}(K_\varepsilon)\geq 1-\varepsilon.$$

La raison d'être de ce théorème est qu'elle permet la stratégie suivante pour démontrer la convergence en loi d'une suite $(X_n)_{n\in\mathbb{N}}$:

1. On démontre que la famille $(X_n)_{n\in\mathbb{N}}$ est tendue;
2. On démontre que la suite $(X_n)_{n\in\mathbb{N}}$ ne peut avoir qu'une unique valeur d'adhérence.

Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires réelles, alors tous les autres modes de convergence cités plus haut, impliquent la convergence en loi.

1. Un espace topologique est dit polonais si c'est un espace métrique complet séparable.

4. Espérances conditionnelles

Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité. Soient maintenant $\mathcal{G}$ une sous-tribu de $\mathcal{F}$ et X une variable aléatoire réelle $\mathcal{F}$ mesurable telle que

$$\mathbb{E}(\mid X \mid) < +\infty.$$

Alors il existe une variable aléatoire Y, $\mathcal{G}$ mesurable telle que:

(1) $\mathbb{E}(\mid Y \mid) < +\infty$;

(2) Pour tout $A \in \mathcal{G}$,

$$\mathbb{E}(1_A X) = \mathbb{E}(1_A Y).$$

De plus, si Z est une autre variable aléatoire $\mathcal{G}$ mesurable qui vérifie les 2 propriétés précédentes, alors

$$\mathbb{P}(Y = Z) = 1.$$

La variable Y s'appelle une version de l'espérance conditionnelle de X sachant $\mathcal{G}$ et on note

$$Y = \mathbb{E}(X \mid \mathcal{G}).$$

Si on considère des variables aléatoires X_1 et X_2 définies sur $(\Omega,\mathcal{F},\mathbb{P})$ telles que

$$\mathbb{E}(\mid X_1 \mid) < +\infty, \mathbb{E}(\mid X_2 \mid) < +\infty,$$

ainsi que deux sous-tribus $\mathcal{H}$ et $\mathcal{G}$ de $\mathcal{F}$ vérifiant:

$$\emptyset \neq \mathcal{H} \subset \mathcal{G},$$

alors on a les propriétés suivantes:

- Si X_1 est $\mathcal{G}$ mesurable, $\mathbb{E}(X_1 \mid \mathcal{G}) = X_1$;
- Si Y est une variable aléatoire $\mathcal{G}$ mesurable positive et bornée alors
$$\mathbb{E}(YX_1) = \mathbb{E}\left(Y\mathbb{E}\left(X_1 \mid \mathcal{G}\right)\right);$$
- Si Y_1 et Y_2 sont deux variables aléatoires positives bornées et $\mathcal{G}$ mesurables, alors
$$\mathbb{E}\left(Y_1X_1 + Y_2X_2 \mid \mathcal{G}\right) = Y_1\mathbb{E}\left(X_1 \mid \mathcal{G}\right) + Y_2\mathbb{E}\left(X_2 \mid \mathcal{G}\right);$$
- $\mathbb{E}\left(\mathbb{E}\left(X_1 \mid \mathcal{G}\right) \mid \mathcal{H}\right) = \mathbb{E}\left(X_1 \mid \mathcal{H}\right)$;
- Si X_n est une suite de variables intégrables qui converge dans $\mathbf{L}^1$ vers X_1, alors
$$\lim_{n\to+\infty} \mathbb{E}\left(X_n \mid \mathcal{G}\right) = \mathbb{E}\left(X_1 \mid \mathcal{G}\right).$$
- Si $X_1 \geq X_2$, $\mathbb{E}\left(X_1 \mid \mathcal{G}\right) \geq \mathbb{E}\left(X_2 \mid \mathcal{G}\right)$;
- Si $\psi : \mathbb{R} \to \mathbb{R}$ est une fonction convexe telle que $\psi(X_1)$ est intégrable alors on a l'inégalité de Jensen:
$$\mathbb{E}(\psi(X_1) \mid \mathcal{G}) \geq \psi\left(\mathbb{E}\left(X_1 \mid \mathcal{G}\right)\right).$$

5. Lois infiniment divisibles et théorème de Lévy-Khinchin

Soient $(\Omega,\mathcal{F},\mathbb{P})$ un espace de probabilité et X une variable aléatoire réelle.

Definition 12. *On dit que X est une variable aléatoire infiniment divisible si pour tout $n \geq 1$, on peut trouver des variables aléatoires indépendantes $X_1,...,X_n$ telles que*

$$X_1 + ... + X_n =^{loi} X.$$

On a alors le théorème important suivant qui caractérise entièrement les variables aléatoires infiniment divisibles:

Théorème 13. *(Lévy-Khinchin) Soit X une variable aléatoire infiniment divisible. On peut trouver $\lambda \in \mathbb{R}$, $\sigma > 0$ et une mesure ν sur $\mathbb{R}\backslash\{0\}$ tels que*

$$\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < +\infty,$$

$$\forall \theta \in \mathbb{R}, \quad \mathbb{E}(e^{i\theta X}) = e^{i\lambda\theta - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}}(e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|\leq 1})\nu(dx)}.$$

Réciproquement, soient $\lambda \in \mathbb{R}$, $\sigma > 0$ et une mesure ν sur $\mathbb{R}\backslash\{0\}$ telle que

$$\int_{\mathbb{R}} (1 \wedge x^2)\nu(dx) < +\infty.$$

La fonction:

$$\theta \to e^{i\lambda\theta - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}}(e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|\leq 1})\nu(dx)}$$

est alors la fonction caractéristique d'une loi infiniment divisible.