

Formulaire de probabilités

Master MIGS/PMG 2025-2026

Yoann Offret

Table des matières

1	Lemmes de Borel–Cantelli	1
2	Convergence en loi	1
3	Loi forte des grands nombres	2
4	Théorème central limite	2
5	Espérance conditionnelle	3
6	Noyaux et probabilités conditionnelles	3
7	Chaînes de Markov	5
7.1	Propriétés générales	5
7.2	Récurrence et transience	6
7.3	Mesure invariante	6
7.4	Théorème ergodique	7
7.5	Théorème de Perron–Frobenius	7
8	Martingales	7

1 Lemmes de Borel–Cantelli

Soit $(A_n)_{n \geq 0}$ une famille d'événements. Alors :

$$\sum_{n \geq 0} \mathbb{P}(A_n) < \infty \implies \mathbb{P}(A_n \text{ infiniment souvent}) = 0.$$

Si les $(A_n)_{n \geq 0}$ sont indépendants et $\sum_{n \geq 0} \mathbb{P}(A_n) = \infty \implies \mathbb{P}(A_n \text{ infiniment souvent}) = 1$.

Ici A_n *infiniment souvent* signifie $\limsup_{n \rightarrow \infty} A_n = \bigcap_{m \geq 1} \bigcup_{n \geq m} A_n$. En particulier, si (A_n) sont indépendants, alors $\mathbb{P}(A_n \text{ i.s.}) = 1$ ssi $\sum_n \mathbb{P}(A_n) = \infty$ (et 0 sinon).

2 Convergence en loi

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires et X une variable aléatoire, toutes deux à valeurs dans $\mathbb{R}^d$. On dit que X_n converge en loi vers X lorsque l'une des propriétés équivalentes suivantes est vérifiée :

1. Pour toute fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ continue et bornée,

$$\mathbb{E}[f(X_n)] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[f(X)].$$

2. Lorsque $d = 1$, pour tout point de continuité t de la fonction de répartition de X ,

$$\mathbb{P}(X_n \leq t) \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(X \leq t).$$

3. (Théorème de Lévy) Pour tout $t \in \mathbb{R}^d$,

$$\mathbb{E}[e^{i\langle t, X_n \rangle}] \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[e^{i\langle t, X \rangle}].$$

On note alors

$$X_n \xrightarrow[n \rightarrow \infty]{} X.$$

3 Loi forte des grands nombres

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires i.i.d. de même loi que X , à valeurs dans un espace mesurable E . Soit $f : E \rightarrow \mathbb{R}$ mesurable telle que $f(X) \in \mathbb{L}^1$. Alors

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \xrightarrow[n \rightarrow \infty]{p.s.} \mathbb{E}[f(X)] = \int_E f(x) \mathbb{P}_X(dx).$$

4 Théorème central limite

Soit $(X_n)_{n \geq 0}$ une suite de v.a. i.i.d. de même loi que X , à valeurs dans un espace mesurable E . Soit $f : E \rightarrow \mathbb{R}$ mesurable telle que $f(X) \in \mathbb{L}^2$. Alors, en posant

$$m = \mathbb{E}[f(X)] \quad \text{et} \quad \sigma^2 = \mathbb{E}[f(X)^2] - (\mathbb{E}[f(X)])^2,$$

on a

$$\sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n f(X_k) - m \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(0, \sigma^2).$$

Version multidimensionnelle : Si $f : E \rightarrow \mathbb{R}^d$ est mesurable et $f(X) \in \mathbb{L}^2$, on pose

$$m = \mathbb{E}[f(X)] \in \mathbb{R}^d, \quad \Gamma = \mathbb{E}[(f(X) - m)(f(X) - m)^\top] \in \mathbb{R}^{d \times d}.$$

Alors

$$\sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n f(X_k) - m \right) \xrightarrow[n \rightarrow \infty]{} \mathcal{N}_d(0, \Gamma),$$

où $\mathcal{N}_d(0, \Gamma)$ désigne la loi gaussienne centrée de matrice de covariance Γ .

5 Espérance conditionnelle

Définition : Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $\mathcal{A} \subset \mathcal{F}$ une sous-tribu et Y une v.a.r.. On dit que $\mathbb{E}[Y|\mathcal{A}]$ est l'espérance conditionnelle de Y sachant $\mathcal{A}$ si :

- $\mathbb{E}[Y|\mathcal{A}]$ est $\mathcal{A}$ -mesurable,
- $\mathbb{E}[Y|\mathcal{A}]$ est $\mathbb{L}^2$, positive ou $\mathbb{L}^1$ selon que $Y \in \mathbb{L}^2$, $Y \geq 0$ ou $Y \in \mathbb{L}^1$ respectivement,
- pour toute v.a. Z respectivement dans $\mathbb{L}^2$, positive, ou bornée,

$$\mathbb{E}[YZ] = \mathbb{E}[\mathbb{E}[Y|\mathcal{A}] Z].$$

Remarques :

- Il suffit de vérifier l'égalité précédente pour $Z = \mathbf{1}_A$, $A \in \mathcal{A}$, ou même pour A appartenant à un π -système engendrant $\mathcal{A}$.
- En particulier, avec $Z = 1$, on retrouve $\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|\mathcal{A}]]$.

Cas particulier : Si X est une v.a. à valeurs dans un espace mesurable E , on note

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

Il existe alors une fonction mesurable $\phi : E \rightarrow \mathbb{R}$ telle que $\mathbb{E}[Y|X] = \phi(X)$ et

$$\mathbb{E}[Yg(X)] = \mathbb{E}[\phi(X)g(X)]$$

pour toute fonction mesurable $g : E \rightarrow \mathbb{R}$ positive ou bornée. La fonction ϕ est unique $\mathbb{P}_X$ -presque partout.

Propriété : Si X, Y sont indépendantes et $H : E \times F \rightarrow \mathbb{R}$ mesurable, alors

$$\mathbb{E}[H(X, Y) | X] = \phi(X), \quad \text{où} \quad \phi(x) = \mathbb{E}[H(x, Y)].$$

Procédure pratique :

1. Geler $X = x$.
2. Calculer $\mathbb{E}[H(x, Y)]$.
3. Dégeler : remplacer x par X .

6 Noyaux et probabilités conditionnelles

Définition : Soient $(E, \mathcal{A})$ et $(F, \mathcal{G})$ deux espaces mesurables. Un *noyau de transition* de E vers F est une application

$$N : E \times \mathcal{G} \longrightarrow [0, \infty]$$

telle que :

1. pour tout $x \in E$, l'application $B \mapsto N(x, B)$ est une mesure sur $(F, \mathcal{G})$,
2. pour tout $B \in \mathcal{G}$, l'application $x \mapsto N(x, B)$ est mesurable sur $(E, \mathcal{A})$.

Autrement dit, c'est une famille de mesures $(N(x, \cdot))_{x \in E}$ qui varie de manière mesurable en x . Si $f : F \rightarrow \mathbb{R}$ est mesurable et $x \in E$, on définit

$$Nf(x) = \int_F f(y) N(x, dy),$$

lorsque cette intégrale est bien définie. On peut montrer que $x \mapsto Nf(x)$ est mesurable. Plus généralement, si $H : E \times F \rightarrow \mathbb{R}$ est mesurable, alors

$$x \mapsto \int_F H(x, y) N(x, dy)$$

est mesurable (quand l'intégrale a un sens).

Définition (loi conditionnelle) : Soient X, Y des v.a. à valeurs respectivement dans E et F . On dit que la loi conditionnelle de Y sachant X est donnée par le noyau N si, pour toute fonction mesurable $f : F \rightarrow \mathbb{R}$ bornée ou positive,

$$\mathbb{E}[f(Y) | X] = Nf(X) = \int_F f(y) N(X, dy).$$

En particulier, on a $N(x, F) = 1 \mathbb{P}_X(dx)$ -presque partout. La loi conditionnelle est alors caractérisée par

$$\mathbb{E}[f(Y)g(X)] = \mathbb{E}[Nf(X)g(X)]$$

pour toutes fonctions muettes f et g appropriées. On dit abusivement que « la loi de Y sachant $X = x$ est $N(x, dy)$ », et on écrit parfois $\mathbb{E}[f(Y) | X = x] = Nf(x)$.

Propriété : Si $H : E \times F \rightarrow \mathbb{R}$ est mesurable et si la loi conditionnelle de Y sachant X est N , alors

$$\mathbb{E}[H(X, Y) | X] = \int_F H(X, y) N(X, dy).$$

Procédure pratique :

1. Geler $X = x$.
2. Calculer $\mathbb{E}[H(x, \tilde{Y})]$ où $\tilde{Y} \sim N(x, dy)$.
3. Dégeler : remplacer x par X .

On en déduit

$$\mathbb{E}[H(X, Y)] = \int_E \left(\int_F H(x, y) N(x, dy) \right) \mathbb{P}_X(dx),$$

c'est-à-dire, de manière concise,

$$\mathbb{P}_{(X, Y)}(dx dy) = \mathbb{P}_X(dx) N(x, dy).$$

7 Chaînes de Markov

7.1 Propriétés générales

Définition : Une suite $(X_n)_{n \geq 0}$ de v.a. à valeurs dans un espace mesurable E (espace d'états) est une *chaîne de Markov de noyau N sur E* si, pour tout $f : E \rightarrow \mathbb{R}$ mesurable positive ou bornée et tout $n \geq 0$, on a

$$\mathbb{E}[f(X_{n+1}) | \mathcal{F}_n] = Nf(X_n),$$

où $(\mathcal{F}_n)_{n \geq 0}$ est une filtration telle que X_n soit $\mathcal{F}_n$ -mesurable. On peut toujours prendre $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$. Le noyau N représente la loi de X_{n+1} sachant $(X_0, \dots, X_n)$, qui ne dépend que de X_n . On note $\mathbb{P}_\nu, \mathbb{E}_\nu$ la probabilité et l'espérance quand $X_0 \sim \nu$.

Cas dénombrable : Si $E = \{e_1, e_2, \dots\}$ est dénombrable, le noyau N s'identifie à une matrice stochastique $P = (P(x, y))_{x, y \in E}$ avec $P(x, y) = N(x, \{y\})$. Alors

$$\mathbb{P}[X_{n+1} = x_{n+1} | X_n = x_n, \dots, X_0 = x_0] = P(x_n, x_{n+1}),$$

dès que $\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) > 0$.

En représentant les mesures comme vecteurs lignes et les fonctions comme vecteurs colonnes :

$$\mathbb{E}_\nu[f(X_n)] = \nu P^n f, \quad \mathbb{E}_x[f(X_n)] = P^n f(x).$$

On a aussi

$$\begin{aligned} \mathbb{P}_\nu(X_0 = x_0, \dots, X_n = x_n) &= \nu(x_0) P(x_0, x_1) \cdots P(x_{n-1}, x_n), \\ \mathbb{P}_\nu(X_n = y) &= \sum_{x_0, \dots, x_{n-1} \in E} \nu(x_0) P(x_0, x_1) \cdots P(x_{n-1}, y). \end{aligned}$$

Définition (irréductibilité) : On dit que la chaîne est *irréductible* si pour tous $x, y \in E$ il existe $n \geq 0$ tel que

$$P^n(x, y) > 0.$$

Autrement dit, à partir de tout état x , il est possible d'atteindre tout état y en un nombre fini d'étapes avec une probabilité positive.

Construction standard : Soient $H : E \times F \rightarrow E$ mesurable, $(U_n)_{n \geq 1}$ i.i.d. de loi sur F , et X_0 indépendante des U_n . En posant

$$X_{n+1} = H(X_n, U_{n+1}),$$

on obtient une chaîne de Markov dont le noyau est la loi de $H(x, U_1)$.

Définition (temps d'arrêt) : Une variable aléatoire $T : \Omega \rightarrow \mathbb{N} \cup \{\infty\}$ est un *temps d'arrêt* relativement à une filtration $(\mathcal{F}_n)_{n \geq 0}$ si, pour tout $n \geq 0$, on a

$$\{T \leq n\} \in \mathcal{F}_n.$$

Autrement dit, au temps n , on sait déjà si l'arrêt est survenu ou non.

Propriété (Markov forte) : Si (X_n) est une chaîne de Markov de noyau N et T un temps d'arrêt alors $(X_{T+n})_{n \geq 0}$ est encore une chaîne de Markov de même noyau (conditionnellement à $T < \infty$). Plus précisément, pour tout $n \geq 0$ et $x \in E$,

$$\mathbb{E}_\nu[f(X_T, \dots, X_{T+n}) | T < \infty, X_T = x] = \mathbb{E}_x[f(X_0, \dots, X_n)].$$

7.2 Récurrence et transience

Définition : Un état $x \in E$ est dit *récurrent* si l'une des conditions équivalentes est satisfaite :

1. $X_n = x$ une infinité de fois p.s. sous $\mathbb{P}_x$.
2. Le temps de retour $T_x = \inf\{n \geq 1 : X_n = x\}$ est fini p.s. sous $\mathbb{P}_x$.
3. La somme $\sum_{k=0}^{\infty} \mathbb{P}_x(X_k = x) = \infty$.

Un état qui n'est pas récurrent est dit *transient*.

Propriété : Si la chaîne est irréductible, alors tous les états sont récurrents ou tous transients.

Exemple : Une chaîne de Markov sur un espace d'états fini possède au moins un état récurrent. Si elle est irréductible, elle est récurrente.

7.3 Mesure invariante

Définition : Une mesure μ sur E est dite *invariante* si

$$\mu N = \mu, \quad \text{c'est-à-dire} \quad \mu N f = \mu f \quad \forall f : E \rightarrow \mathbb{R} \text{ mesurable positive (ou bornée)}.$$

Si $\mu(E) = 1$, on dit que μ est une *probabilité invariante*.

Caractérisation : μ est invariante $\iff (X_0 \sim \mu \implies X_n \sim \mu, \forall n \geq 0)$.

Cas dénombrable : Si P est la matrice de transition, alors μ est invariante si et seulement si $\mu P = \mu$.

Mesure réversible : μ est dite réversible si

$$\mu(x)P(x, y) = \mu(y)P(y, x) \quad \forall x, y \in E.$$

Toute mesure réversible est invariante.

Résultats importants :

- Si la chaîne est irréductible et récurrente, elle possède une unique mesure invariante non triviale (à un facteur près).
- Dans ce cas :

$$\mu(y) = \mathbb{E}_x \left[\sum_{n=0}^{T_x-1} \mathbf{1}_{\{X_n=y\}} \right],$$

où x est un état récurrent arbitraire.

- Dichotomie : $\mu(E) < \infty \implies$ récurrence positive (existence d'une unique probabilité invariante). $\mu(E) = \infty \implies$ récurrence nulle.
- Si la chaîne est irréductible et admet une probabilité invariante, alors elle est récurrente positive.
- Dans le cas récurrent positif, pour tout état x ,

$$\mu(x) = \frac{1}{\mathbb{E}_x[T_x]}.$$

- Si la chaîne est récurrente nulle, alors $\mathbb{E}_x[T_x] = \infty$.

7.4 Théorème ergodique

Si (X_n) est irréductible récurrente positive et μ sa probabilité invariante, alors pour toute loi initiale ν et tout $f \in \mathbb{L}^1(\mu)$:

$$\frac{1}{n} \sum_{k=1}^n f(X_k) \xrightarrow[n \rightarrow \infty]{\mathbb{P}_{\nu\text{-p.s.}}} \int_E f d\mu.$$

En particulier, pour tout $x \in E$,

$$\frac{1}{n} \text{card}\{k \leq n-1 : X_k = x\} \xrightarrow[n \rightarrow \infty]{\mathbb{P}_{\nu\text{-p.s.}}} \mu(x).$$

Si la chaîne est récurrente nulle, les convergences ont lieu vers 0.

7.5 Théorème de Perron–Frobenius

Définition : La période d'un état x est

$$d(x) = \text{pgcd}\{n \geq 1 : P^n(x, x) > 0\}.$$

Propriété : Si la chaîne est irréductible, $d(x)$ ne dépend pas de x ; on parle de période de la chaîne. Si $d = 1$, la chaîne est dite *apériodique*.

Résultat : Soit (X_n) une chaîne irréductible, apériodique et récurrente positive de matrice de transition P et de probabilité invariante μ . Alors, pour toute loi initiale ν ,

$$X_n \xrightarrow[n \rightarrow \infty]{} \mu, \quad \text{c.-à-d.} \quad \mathbb{P}_{\nu}(X_n = x) \xrightarrow[n \rightarrow \infty]{} \mu(x).$$

Dans le cas fini, ce résultat s'obtient en étudiant le spectre de P , c'est un théorème d'algèbre linéaire.

8 Martingales

Définition : Une suite de v.a.r. $(X_n)_{n \geq 0}$ est une $(\mathcal{F}_n)_{n \geq 0}$ -martingale, où $(\mathcal{F}_n)_{n \geq 0}$ est une filtration, si :

1. X_n est $\mathcal{F}_n$ -mesurable ;
2. $X_n \in \mathbb{L}^1$;
3. $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n$.

Si dans 3. on remplace $=$ par $\leq$ (resp. $\geq$), on parle de *sur-martingale* (resp. *sous-martingale*). Par défaut, on considère la filtration canonique $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$.

Propriété : Si (X_n) est une martingale et T un temps d'arrêt, alors $(X_{T \wedge n})_{n \geq 0} \mathbf{1}_{\{T < \infty\}}$ est encore une martingale. De même pour les sur- et sous-martingales. On note $a \wedge b$ le minimum de a et b .

Théorème de convergence presque sûre : Si (X_n) est une martingale (ou sur/sous-martingale) bornée dans $\mathbb{L}^1$, c'est-à-dire

$$\sup_{n \geq 0} \mathbb{E}[|X_n|] < \infty,$$

alors il existe $X_\infty \in \mathbb{L}^1$ telle que

$$X_n \xrightarrow[n \rightarrow \infty]{p.s.} X_\infty.$$

Cette hypothèse est satisfaite notamment si :

1. (X_n) est bornée ;
2. (X_n) est une sur-martingale positive ;
3. (X_n) est une sous-martingale majorée par une variable intégrable.

Théorème de convergence dans $\mathbb{L}^2$: Si (X_n) est une martingale bornée dans $\mathbb{L}^2$, c'est-à-dire

$$\sup_{n \geq 0} \mathbb{E}[|X_n|^2] < \infty,$$

alors il existe $X_\infty \in \mathbb{L}^2$ telle que

$$X_n \xrightarrow[n \rightarrow \infty]{p.s.} X_\infty \quad \text{et} \quad X_n \xrightarrow[n \rightarrow \infty]{\mathbb{L}^2} X_\infty.$$

De plus,

$$X_n = \mathbb{E}[X_\infty | \mathcal{F}_n].$$