Cours de Convergence, Intégration, Probabilités

Mesures sur un espace produit. Indépendance de variables aléatoires

Séance 8 - Convolution et densité des espaces L^p
Probabilités dans $\mathbb{R}^N$. Indépendance

CentraleSupélec - Cursus ingénieur

17 octobre 2019

Amphis CIP 6, 7, 8 et 9

- Hervé MOUTARDE
  Institut de recherche sur les lois fondamentales de l'univers (IRFU), CEA, Université Paris-Saclay
  Orme des Merisiers, Bât. 703
  `herve.moutarde@cea.fr`

Des questions ?

- daskit.com/cip19-20 puis section "Amphi 8".

Support

- Support amphi 8 en version vierge disponible dès à présent sur edunao.
- Support amphi 8 en version annotée disponible ultérieurement.

Quelques éléments des CM et TD précédents

- Tribu (**def. 3.3**), tribu des événements (**def. 3.25**) et tribu produit (**def. VII.1.1**).
- Mesure (**def. 3.21**), mesure de probabilité (**def. 3.26**) et mesure produit (**th. VII.1.4**).
- Changement de variables (**th. VII.2.3**).
- Ordre d'intégration dans les intégrales multiples et fonctions sommables (**th. VII.2.4** et **th. VII.2.5**).
- Indépendance de variables aléatoires (**def. VII.1.5**).
- Indépendance de variables aléatoires et mesure produit (**prop. VII.1.6**).
- Fonction caractéristique (**def. VII.2.6**).
- Indépendance de variables aléatoires : expression en termes de fonctions caractéristiques (**VII.2.7**).

Programme

Objectifs de la séance

- Je suis capable d'étudier la **convolution** de deux fonctions (intégrabilité, convergence).
- Je suis capable d'exprimer l'**indépendance** de variables aléatoires en terme de **mesure produit**.
- Je suis capable de **vérifier** que deux variables aléatoires sont **indépendantes**.
- Je suis capable de **déterminer la loi** d'une variable aléatoire définie comme fonction de deux variables aléatoires indépendantes.
- Je distingue parfaitement les notions d'**indépendance** et de **non corrélation**.
- Je suis capable d'étudier un **vecteur de variables aléatoires réelles**, dont la loi est donnée (loi de chaque composante, indépendance).

Produit de convolution

Soient f et g deux fonctions mesurables sur $\mathbb{R}^N$.

On considère $\forall x \in \mathbb{R}^N, \quad f * g(x) = \int_{\mathbb{R}^N} f(x-y)g(y) \ \lambda(dy)$.

Théorème VIII.1.1

Soient $f, g \in L^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda)$.
Alors pour λ-presque tout $x \in \mathbb{R}^N$, $f * g(x)$ est bien définie.
De plus, $f * g \in L^1$ et $\|f * g\|_1 \leq \|f\|_1 \ \|g\|_1$.

Proposition VIII.1.2 (Admis)

Soient $p \in [1, +\infty[$ et $q \in]1, +\infty[$ tels que $\frac{1}{p} + \frac{1}{q} = 1$.
Soient $f \in L^p$ et $g \in L^q$. Alors pour $x \in \mathbb{R}^N$, $f * g(x)$ est bien définie et la fonction $f * g$ est uniformément continue et bornée sur $\mathbb{R}^N$.

Approximation de la mesure de Dirac

Définition VIII.1.3 (Suite régularisante)

Une suite $(\varphi_n)_{n\in\mathbb{N}}$ *dans* $C_c(\mathbb{R}^N)$ *(fonctions continues à support compact) est une approximation de* δ_0 *si :*

- *Il existe un compact* K *tel que supp* $\varphi_n \subset K$ *pour tout* n
- $\forall n,\ \varphi_n \geq 0$ *et* $\displaystyle\int_{\mathbb{R}^N} \varphi_n . d\lambda = 1.$
- $\forall \delta > 0,\ \lim_{n\to\infty} \displaystyle\int_{\{|x|>\delta\}} \varphi_n(x)\ \lambda(dx) = 0.$

Approximation de la mesure de Dirac

Définition VIII.1.3 (Suite régularisante)

Une suite $(\varphi_n)_{n\in\mathbb{N}}$ dans $C_c(\mathbb{R}^N)$ (fonctions continues à support compact) est une approximation de δ_0 si :

- *Il existe un compact K tel que supp $\varphi_n \subset K$ pour tout n*
- $\forall n$, $\varphi_n \geq 0$ *et* $\displaystyle\int_{\mathbb{R}^N} \varphi_n . d\lambda = 1$.
- $\forall \delta > 0$, $\displaystyle\lim_{n\to\infty} \int_{\{|x|>\delta\}} \varphi_n(x)\ \lambda(dx) = 0$.

Exemple : Si $\varphi : \mathbb{R}^N \to \mathbb{R}$ est continue à support compact tel que $\int \varphi d\lambda = 1$, alors on pose : $\forall n, \quad \varphi_n(x) = n^N\ \varphi(nx)$.

Densité dans L^p

Proposition VIII.1.4 (Partiellement admis)

Soit $(\varphi_n)_{n\in\mathbb{N}}$ une approximation de δ_0.

(i) *Si $f:\mathbb{R}^N\to\mathbb{R}$ est continue, on a $\varphi_n * f \to f$ uniformément sur tout compact.*

(ii) *Si $f\in L^p$, avec $p\in[1,+\infty[$, on a $\varphi_n * f\to f$ dans L^p.*

Densité dans L^p

Théorème VIII.1.5

Pour Ω ouvert connexe de $\mathbb{R}^N$, l'ensemble $\mathcal{D}(\Omega) = C_c^\infty(\Omega)$ est dense dans $L^p(\Omega, \lambda)$ pour $p \in [1, +\infty[$.

Mesure de probabilité sur $\Omega = \mathbb{R}^N$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^N)$

La tribu $\boxed{\mathcal{B}(\mathbb{R}^N) \text{ est trop grande}}$ pour déterminer la mesure $\mathbf{P}$ de manière directe (même cas que $\mathbb{R}$).

Mesure de probabilité sur $\Omega = \mathbb{R}^N$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^N)$

La tribu $\boxed{\mathcal{B}(\mathbb{R}^N) \text{ est trop grande}}$ pour déterminer la mesure $\mathbf{P}$ de manière directe (même cas que $\mathbb{R}$).

On utilise alors la définition de $\mathcal{B}(\mathbb{R}^N)$ comme tribu engendrée par $\{]-\infty, x_1] \times \cdots \times]-\infty, x_N];\ x_1, \ldots, x_N \in \mathbb{R}\} \subset \Omega$.

Mesure de probabilité sur $\Omega = \mathbb{R}^N$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^N)$

La tribu $\boxed{\mathcal{B}(\mathbb{R}^N) \text{ est trop grande}}$ pour déterminer la mesure $\mathbf{P}$ de manière directe (même cas que $\mathbb{R}$).

On utilise alors la définition de $\mathcal{B}(\mathbb{R}^N)$ comme tribu engendrée par $\{]-\infty, x_1] \times \cdots \times]-\infty, x_N];\ x_1, \ldots, x_N \in \mathbb{R}\} \subset \Omega$.

- π-système
- La mesure $\mathbf{P}$ est caractérisée par

$$(x_1, \ldots, x_N) \mapsto \mathbf{P}(]-\infty, x_1] \times \cdots \times]-\infty, x_N]) = F(x_1, \ldots, x_N)$$

Mesure de probabilité sur $\Omega = \mathbb{R}^N$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^N)$

La tribu $\boxed{\mathcal{B}(\mathbb{R}^N) \text{ est trop grande}}$ pour déterminer la mesure $\mathbf{P}$ de manière directe (même cas que $\mathbb{R}$).

On utilise alors la définition de $\mathcal{B}(\mathbb{R}^N)$ comme tribu engendrée par $\{]-\infty, x_1] \times \cdots \times]-\infty, x_N];\ x_1, \ldots, x_N \in \mathbb{R}\} \subset \Omega$.

- π-système
- La mesure $\mathbf{P}$ est caractérisée par

$$(x_1, \ldots, x_N) \mapsto \mathbf{P}(]-\infty, x_1] \times \cdots \times]-\infty, x_N]) = F(x_1, \ldots, x_N)$$

La fonction de répartition dans $\mathbb{R}^N$ est moins utilisée que dans $\mathbb{R}$.

Définition VIII.2.1

Soit $X = (X_1, \ldots, X_N)$ *un vecteur aléatoire.*
Si $X_k \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$ *(pour tout* $k = 1, \ldots, N$*), alors le vecteur*

$$\mathbf{E}[X] = (\mathbf{E}[X_1], \ldots, \mathbf{E}[X_N])$$

est appelé **espérance de** X*.*

Définition VIII.2.1

Soit $X = (X_1, \ldots, X_N)$ *un vecteur aléatoire.*
Si $X_k \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$ *(pour tout* $k = 1, \ldots, N$*), alors le vecteur*

$$\mathbf{E}[X] = (\mathbf{E}[X_1], \ldots, \mathbf{E}[X_N])$$

est appelé **espérance de** X.

Définition VIII.2.2

Soient X *et* Y *deux variables aléatoires réelles dans* $\mathbf{L}^2(\Omega, \mathcal{F}, \mathbf{P})$*.*
On définit la **covariance de** X **et** Y *comme la quantité*

$$\mathrm{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y].$$

Proposition VIII.2.3

Si X et Y sont des variables aléatoires réelles admettant des moments d'ordre 2, alors

(i) *L'application*

$$\begin{array}{ccc} \mathbf{L}^2(\Omega,\mathcal{F},\mathbf{P})\times\mathbf{L}^2(\Omega,\mathcal{F},\mathbf{P}) & \to & \mathbb{R} \\ (X,Y) & \mapsto & \mathrm{Cov}(X,Y) \end{array}$$

est bilinéaire.

(ii) $\mathrm{Cov}(X,X)=\mathrm{Var}(X)$.

(iii) $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$.

(iv) $\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\,\mathrm{Cov}(X,Y)$.

Définition VIII.2.4 (Corrélation entre 2 v.a.)

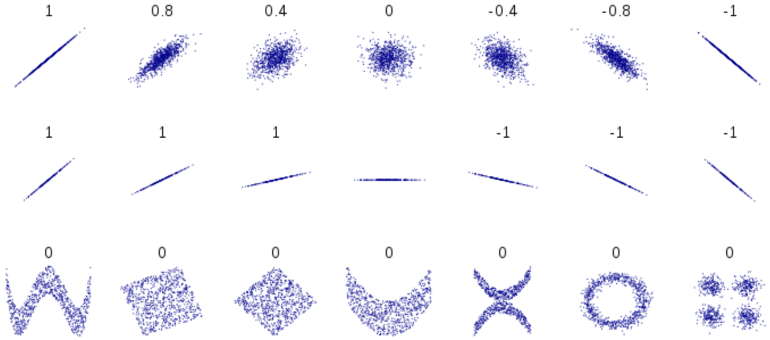
Soient X et Y sont deux variables aléatoires réelles admettant des moments d'ordre 2 non nuls. La quantité

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

où $\sigma(X) = \sqrt{\mathrm{Var}(X)}$ et $\sigma(Y) = \sqrt{\mathrm{Var}(Y)}$ sont les écart-types de X et Y, est appelée **coefficient de corrélation entre X et Y**.

Attention à ne pas mal interpréter la notion de corrélation !

Corrélation (statistiques) — Wikipédia, l'encyclopédie libre



Attention à ne pas confondre corrélation et causalité !

Définition VIII.2.5 (Matrice de covariances)

Soit $X = (X_1, \dots, X_N)$ *un vecteur aléatoire dont chaque composante admet un moment d'ordre* 2. *La* **matrice de covariances** *de* X *est la matrice* $\Sigma = [c_{ij}]_{1 \leq i,j \leq N}$ *telle que*

$$\forall i, j = 1, \dots, N, \quad c_{ij} = \mathrm{Cov}(X_i, X_j).$$

Proposition VIII.2.6

Toute matrice de covariances $\Sigma = [c_{ij}]_{1 \leq i,j \leq N}$ *est* **symétrique** *et* **positive**, *c'est-à-dire*

- *Pour tous* $i, j = 1, \ldots, N$, *on a* $c_{ij} = c_{ji}$.
- *Pour tout vecteur* $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{R}^N$, *on a*

$$\lambda \Sigma \lambda^t = \sum_{1 \leq i,j \leq N} \lambda_i c_{ij} \lambda_j \geq 0.$$

Théorème VIII.2.7

Soit $X = (X_1, \ldots, X_N)$ *un vecteur aléatoire admettant une densité de probabilité* f_X*.*
Si $h : \Delta \subset \mathbb{R}^N \to D \subset \mathbb{R}^N$ *est un* C^1*-difféomorphisme, alors le vecteur aléatoire* $Y = h(X)$ *admet la densité de probabilité* f_Y *définie par*

$$\forall y \in \mathbb{R}^N, \quad f_Y(y) = \frac{f_X(h^{-1}(y))}{|\det(Jh(h^{-1}(y)))|} \mathbb{1}_D(y)$$

où Jh *est la matrice jacobienne de* h*.*

Définition VIII.2.8

Soit $X = (X_1, \ldots, X_N)$ un vecteur aléatoire. La loi d'un sous-vecteur $(X_{i_1}, \ldots, X_{i_k})$ est appelée **loi marginale** *de la mesure de probabilité P_X.*
En particulier, les lois des variables aléatoires réelles X_i sont des lois marginales de P_X.

Proposition VIII.2.9

Soit $X = (X_1, \dots, X_N)$ *de fonction de répartition*

$$F_X(x_1, \dots, x_N) = \mathbf{P}(X_1 \leq x_1, \dots, X_N \leq x_N).$$

$\forall 1 \leq k < N$, *la fonction de répartition de* $(X_1, \dots, X_k)$ *est*

$$\begin{aligned} F_{(X_1,\dots,X_k)}(x_1, \dots, x_k) &= \lim_{x_{k+1},\dots,x_N \to +\infty} F_X(x_1, \dots, x_N) \\ &= F_X(x_1, \dots, x_k, +\infty, \dots, +\infty). \end{aligned}$$

Proposition VIII.2.10

Soit $X = (X_1, \ldots, X_N)$ admettant une densité de probabilité f_X.
Le vecteur aléatoire $(X_1, \ldots, X_k)$ admet la densité définie par

$$f_{(X_1,\ldots,X_k)}(x_1, \ldots, x_k) = \int_{\mathbb{R}^{N-k}} f_X(x_1, \ldots, x_N)\ \lambda(dx_{k+1} \ldots dx_N).$$

Définition VIII.3.1 (Rappel, tribus indépendantes)

Dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$*, des sous-tribus* $(\mathcal{F}_i)_{i \in I}$ *de* $\mathcal{F}$ *sont dites* **indépendantes** *si pour tout sous-ensemble fini* J *de* I *et toute famille d'événements* $A_i \in \mathcal{F}_i$ *(avec* $i \in I$*),*

$$\mathbf{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbf{P}(A_i).$$

Définition VIII.3.2 (Variables aléatoires indépendantes)

Soit $(X_i)_{i \in I}$ *une famille de variables aléatoires, à valeurs dans les espaces* $(E_i, \mathcal{E}_i)$*. Les variables* X_i *sont dites* **indépendantes** *si les sous-tribus engendrées* $X_i^{-1}(\mathcal{E}_i)$ *sont indépendantes.*

Théorème VIII.3.3

Deux v.a. $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ et $Y : (\Omega, \mathcal{F}) \to (\tilde{E}, \tilde{\mathcal{E}})$ sont indépendantes si l'une des conditions équivalentes suivantes est vérifiée :

(a) *Pour tous $A \in \mathcal{E}$ et $B \in \tilde{\mathcal{E}}$,*

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\ \mathbf{P}(Y \in B);$$

(b) *Pour toutes fonctions mesurables bornées $f : E \to \mathbb{R}$ et $g : \tilde{E} \to \mathbb{R}$,*

$$\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\ \mathbf{E}[g(Y)];$$

(c) *Pour toutes fonctions mesurables bornées $f : E \to \mathbb{R}$ et $g : \tilde{E} \to \mathbb{R}$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.*

(a) Pour tous $A \in \mathcal{E}$ et $B \in \tilde{\mathcal{E}}$,

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A)\ \mathbf{P}(Y \in B);$$

(a) Pour toutes fonctions mesurables bornées $f : E \to \mathbb{R}$ et $g : \tilde{E} \to \mathbb{R}$,

$$\mathbf{E}[f(X)g(Y)] = \mathbf{E}[f(X)]\ \mathbf{E}[g(Y)];$$

(i) Pour toutes fonctions mesurables bornées $f : E \to \mathbb{R}$ et $g : \tilde{E} \to \mathbb{R}$, les variables aléatoires $f(X)$ et $g(Y)$ sont indépendantes.

Proposition VIII.3.9

Soient X et Y deux variables aléatoires réelles dans $\mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$. Si X et Y sont indépendantes, alors on a

$$\mathbf{E}[XY] = \mathbf{E}[X]\ \mathbf{E}[Y].$$

Si X et Y sont dans $\mathbf{L}^2(\Omega, \mathcal{F}, \mathbf{P})$ et sont indépendantes, alors elles sont non-corrélées, i. e. $\mathrm{Cov}(X, Y) = 0$ ce qui s'écrit encore

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Définition VII.2.6 (Rappel, fonction caractéristique)

On appelle **fonction caractéristique** d'une variable aléatoire X à valeurs dans $\mathbb{R}^N$ l'application $\varphi_X : \mathbb{R}^N \to \mathbb{C}$ définie par

$$t \mapsto \varphi_X(t) = \mathbf{E}\left[e^{i\langle t,X\rangle}\right] = \int_{\mathbb{R}^N} e^{i\langle t,x\rangle} P_X(dx)$$

$$t \mapsto \varphi_X(t) = \mathbf{E}\left[e^{i\langle t,X\rangle}\right] = \int_{\mathbb{R}^N} e^{i\langle t,x\rangle} P_X(dx)$$

Proposition VIII.4.1

- $\varphi_X(0) = 1$ *et pour tout* $t \in \mathbb{R}^N$, $|\varphi_X(t)| \leq 1$.
- *Pour tout* $\lambda \in \mathbb{R}$ *et tout* $a \in \mathbb{R}^N$, $\varphi_{\lambda X+a}(t) = e^{iat}\varphi_X(\lambda t)$.
- φ_X *est une fonction semi-positive, i.e. pour tous* $n \geq 1$ *et tous* $t_1, \ldots, t_n \in \mathbb{R}^N$, *on a*

$$\forall z_1, \ldots, z_n \in \mathbb{C}, \qquad \sum_{1\leq j,k\leq n} z_j\, \varphi_X(t_j - t_k)\, \overline{z_k} \geq 0.$$

Proposition VIII.4.2

La fonction caractéristique d'une v.a. X à valeurs dans $\mathbb{R}^N$ est continue sur $\mathbb{R}^N$.

Proposition VIII.4.3

Si la loi de X admet une densité de probabilité, alors

$$\lim_{|t|\to\infty} \varphi_X(t) = 0.$$

Théorème VIII.4.4 (Admis, théorème d'inversion)

Si la fonction caractéristique φ_X d'une v.a. X est dans $\mathbf{L}^1(\mathbb{R})$, alors X admet la densité $f_X : \mathbb{R}^N \to \mathbb{R}_+$ définie par

$$\forall x \in \mathbb{R}^N, \quad f_X(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{-i\langle t,x\rangle} \varphi_X(t) \ \lambda(dt).$$

Théorème VIII.4.4 (Admis, théorème d'inversion)

Si la fonction caractéristique φ_X d'une v.a. X est dans $\mathbf{L}^1(\mathbb{R})$, alors X admet la densité $f_X : \mathbb{R}^N \to \mathbb{R}_+$ définie par

$$\forall x \in \mathbb{R}^N, \quad f_X(x) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} e^{-i\langle t,x\rangle} \varphi_X(t) \ \lambda(dt).$$

Théorème VIII.4.5 (Théorème d'unicité)

Deux variables aléatoires X et Y ont même loi si et seulement si $\varphi_X = \varphi_Y$.

Théorème VII.2.7 (Rappel)

Les v.a. réelles $X_1, \ldots, X_N$ sont indépendantes si et seulement si leurs fonctions caractéristiques vérifient

$$\forall t \in \mathbb{R}^N, \quad \varphi_X(t) = \prod_{k=1}^{N} \varphi_{X_k}(t_k)$$

où $X = (X_1, \ldots, X_N)$.

Proposition VIII.4.6

Soient $X_1, \dots, X_n$ des vecteurs aléatoires **indépendants** *à valeurs dans $\mathbb{R}^N$, de lois respectives $P_{X_1}, \dots, P_{X_n}$.*
La loi de $X_1 + \cdots + X_n$ est le produit de convolution $P_{X_1} * \cdots * P_{X_n}$ et a pour fonction caractéristique $\varphi_{X_1+\cdots+X_n} = \prod_{j=1}^{n} \varphi_{X_j}$.

Proposition VIII.4.7

Soit X une v.a. à valeurs réelles dans $\mathbf{L}^n(\Omega, \mathcal{F}, \mathbf{P})$ ($n \geq 1$). Alors, sa fonction caractéristique φ_X est de classe C^n et

$$\forall t \in \mathbb{R}, \quad \varphi_X^{(n)}(t) = i^n \, \mathbf{E}[X^n e^{itX}].$$

En particulier, $\mathbf{E}[X^n] = i^{-n} \varphi_X^{(n)}(0)$.

Objectifs de la séance

- Je suis capable d'étudier la **convolution** de deux fonctions (**def. VIII.1.1** et **prop. VIII.1.4**).
- Je suis capable d'exprimer l'**indépendance** de variables aléatoires en terme de **mesure produit** (**prop.VII.1.6**).
- Je suis capable de **vérifier** que deux variables aléatoires sont **indépendantes** (**th.VIII.3.3)**.
- Je suis capable de **déterminer la loi** d'une variable aléatoire définie comme fonction de deux variables aléatoires indépendantes (**prop.VII.1.7** et **prop.VII.1.8**).
- Je distingue parfaitement les notions d'**indépendance** (**def. VIII.3.1**) et de **non corrélation** (**VIII.2.4**).
- Je suis capable d'étudier un **vecteur de variables aléatoires réelles** (**def. VIII.2.1**), dont la loi est donnée (loi de chaque composante (**def. VIII.2.8**), indépendance (**th. VIII.3.3** et **th. VII.2.7**).

Références bibliographiques

- T. Gallouët, R. Herbin. *Mesure, intégration, probabilités*.
  https://www.i2m.univ-amu.fr/perso/thierry.gallouet/licence.d/mes-int-pro.pdf
- O. Garet. *Intégration et probabilités*.
  http://www.iecl.univ-lorraine.fr/~Olivier.Garet/cours/ip/ip-2016-2017.pdf
- J.-F. Le Gall. *Intégration, probabilités et processus aléatoires*.
  https://www.math.u-psud.fr/~jflegall/IPPA2.pdf
- W. Rudin. *Real and Complex Analysis*. McGraw-Hill.