Cours de Convergence, Intégration, Probabilités et Equations aux Dérivées Partielles

Mesures et intégrales. Intégrale de Lebesgue. Espaces de probabilité

Séance 6 - Probabilités sur $\mathbb{R}$, variables aléatoires réelles

CentraleSupélec - Cursus ingénieur

3 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`

- Analyse théorique de données expérimentales :
  - Interaction forte, récemment ondes gravitationnelles.
  - Modélisation (structure du proton).
  - Analyse statistique de données.
  - Développement de codes de calcul scientifique.

- Dans le cursus CentraleSupelec :
  - 1A CIP et EDP.
  - 3A Théorie quantique du champ.

Des questions ?

- daskit.com/cip19-20 puis section "Amphi 6".

Délégués de cours de CIP modalité normal :

- Alix CHAZOTTES,
- Laure COQUELET,
- Simon MARTEL,
- Guillaume RIPERT,
- Damien TASSO.

Support

- Support amphi 6 en version vierge disponible dès à présent sur edunao.
- Support amphi 6 en version annotée disponible ultérieurement.
- Enregistrement vidéo disponible sur la web tv après validation.

Quelques éléments des CM et TD précédents

- Tribu (**def. 3.3**) et tribu des événements (**def. 3.25**).
- Mesure (**def. 3.21**) et mesure de probabilité (**def. 3.26**).
- Fonction mesurable (**def. 3.10**) et variable aléatoire *discrète* (**def. 3.38**).
- Loi d'une variable aléatoire *discrète* (**def. 3.39**).
- Formalisme homogène pour les cas discrets et continus (**TD Exercice IV.1**).
- $\mathcal{B}(\mathbb{R}) = \sigma(\{]-\infty, a[/a \in \mathbb{R}\})$ (**TD Exercice III.2**).
- Construction de l'intégrale relativement à une mesure μ.
- Espaces $L^p(\Omega)$ (**prop. 4.20**).

Programme

1. Construction d'une mesure de probabilité
   - Exemples de mesure de probabilité
   - Mesure de probabilité sur $\mathbb{R}$
2. Variables aléatoires
   - Loi d'une variable aléatoire
   - Intégration, moments
   - Moments d'ordre $n > 1$
3. Exemples de lois de probabilité

Objectifs de la séance

- je maîtrise les notions d'**espace de probabilité**, de **variable aléatoire**, de **loi** ;
- je suis capable de calculer la **probabilité d'un événement**, lorsque la mesure de probabilité est donnée ;
- je maîtrise les notions de **fonction de répartition** et de **densité de probabilité** ;
- je sais déterminer la loi d'une variable aléatoire ;
- je suis capable de vérifier qu'une variable aléatoire donnée est mesurable par rapport à une sous-tribu ;
- je suis capable de calculer l'**espérance** et la **variance** d'une variable aléatoire, lorsqu'elles existent ;
- je maîtrise l'application du **théorème de transfert** (calculs, détermination de lois).

Deux exemples de mesure de probabilité sur $(\Omega, \mathcal{F})$

1. *Mesure discrète* de support $\{a_n;\ n \in \mathbb{N}\}$: $\mathbf{P} = \sum_{n=0}^{\infty} \alpha_n\ \delta_{a_n}$, où
   - pour tout $n \in \mathbb{N}$, δ_{a_n} est la mesure de Dirac en a_n ;
   - $\forall n, \alpha_n \geq 0$ et $\sum_{n=0}^{\infty} \alpha_n = 1$.
2. *La mesure* $\mathbf{P} = f.\mu$, où
   - μ est une mesure sur $(\Omega, \mathcal{F})$
   - et $f : (\Omega, \mathcal{F}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ une application mesurable, intégrable par rapport à μ et telle que $\int_\Omega f(x)\ \mu(dx) = 1$,

   définie par

$$\forall A \in \mathcal{F}, \quad \mathbf{P}(A) = \int_A f(x)\ \mu(dx) := \int_\Omega f(x)\ \mathbf{1}_A(x)\ \mu(dx).$$

En général : pour un espace d'état Ω non dénombrable

La définition d'une mesure de probabilité sur une tribu $\mathcal{F}$ est impossible de manière directe (si $\mathcal{F}$ est trop grande).

En général : pour un espace d'état Ω non dénombrable

La définition d'une mesure de probabilité sur une tribu $\mathcal{F}$ est impossible de manière directe (si $\mathcal{F}$ est trop grande).

On utilise alors la définition de $\mathcal{F}$ comme tribu engendrée par une partie $\mathcal{I} \subset \mathcal{P}(\Omega)$.

- Comment définir une mesure de probabilité $\mathbf{P} : \mathcal{F} \to [0, 1]$, à partir des valeurs sur $\mathcal{I}$?
- Existence et unicité de l'extension de $\mathcal{I}$ à $\mathcal{F}$?
- Conditions sur $\mathcal{I}$: algèbre de Boole et π-système.

Définition VI.1.1 (π-système)

*Une classe $\mathcal{I}$ de sous-ensembles de Ω est appelée π-**système** sur Ω si elle est stable par intersection finie.*

Lemme VI.1.2 (Admis)

Si deux mesures de probabilité coincident sur un π-système $\mathcal{I}$, alors elles coincident aussi sur la tribu $\sigma(\mathcal{I})$ engendrée par $\mathcal{I}$.

Définition VI.1.3 (Algèbre de Boole)

Une famille de parties de Ω est appelée **algèbre de Boole** *si :*

(i) *Elle contient Ω.*
(ii) *Elle est stable par passage au complémentaire.*
(iii) *Elle est stable par union finie.*

Théorème VI.1.4 (Carathéodory, admis)

Soit $\mathcal{F}_0$ une algèbre de Boole sur Ω et $\mathcal{F} = \sigma(\mathcal{F}_0)$.

- *Si $\mu_0 : \mathcal{F}_0 \to [0, +\infty]$ est une fonction σ-additive, alors il existe une mesure μ sur $(\Omega, \mathcal{F})$ telle que $\mu = \mu_0$ sur $\mathcal{F}_0$.*
- *Si de plus $\mu_0(\Omega) < +\infty$, alors cette extension est unique.*

Fonction de répartition

Définition VI.1.5 (Fonction de répartition)

Dans l'espace de probabilité $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P})$, *où* $\mathcal{B}(\mathbb{R})$ *est la tribu de Borel de* $\mathbb{R}$, *la* **fonction de répartition** *de* $\mathbf{P}$ *est l'application*

$$\begin{aligned} F : \mathbb{R} &\to [0,1] \\ x &\mapsto F(x) = \mathbf{P}(]-\infty, x]). \end{aligned}$$

Fonction de répartition

Définition VI.1.5 (Fonction de répartition)

Dans l'espace de probabilité $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P})$, *où* $\mathcal{B}(\mathbb{R})$ *est la tribu de Borel de* $\mathbb{R}$, *la* **fonction de répartition** *de* $\mathbf{P}$ *est l'application*

$$\begin{aligned} F : \mathbb{R} &\to [0,1] \\ x &\mapsto F(x) = \mathbf{P}(]-\infty, x]). \end{aligned}$$

Théorème VI.1.6

Sur l'espace de probabilité $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P})$, *la mesure de probabilité* $\mathbf{P}$ *est caractérisée par sa fonction de répartition* F.

Fonction de répartition

Théorème VI.1.7 (Partiellement admis)

Une fonction F est la fonction de répartition d'une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ si et seulement si les 3 conditions suivantes sont vérifiées :

(i) *F est croissante ;*

(ii) *F est continue à droite ;*

(iii) $\lim_{x \to -\infty} F(x) = 0$ *et* $\lim_{x \to +\infty} F(x) = 1$.

Fonction de répartition

Proposition VI.1.8

Soit F la fonction de répartition d'une mesure de probabilité $\mathbf{P}$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Pour tout réel x, on a

$$\mathbf{P}(\{x\}) = F(x) - F(x-).$$

Cas particulier : densité de probabilité

Définition VI.1.9 (Densité de probabilité)

Si une application $f : \mathbb{R} \to \mathbb{R}$ est positive, intégrable par rapport à la mesure de Lebesgue λ et $\int_{\mathbb{R}} f(x).\lambda(dx) = 1$, alors l'application

$$F : x \mapsto F(x) = \int_{]-\infty,x]} f(t).\lambda(dt)$$

est la fonction de répartition d'une mesure de probabilité $\mathbf{P}$ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
La fonction f est appelée **densité** *de $\mathbf{P}$.*
$\mathbf{P}$ est dite **absolument continue** *par rapport à la mesure de Lebesgue λ.*

Cas particulier : densité de probabilité

$\mathbf{P}$ est définie par

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad \mathbf{P}(A) = \int_A f(x)\ \lambda(dx) = \int_\Omega f(x)\ \mathbf{1}_A(x)\ \lambda(dx).$$

Cas particulier : densité de probabilité

$\mathbf{P}$ est définie par

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad \mathbf{P}(A) = \int_A f(x)\ \lambda(dx) = \int_\Omega f(x)\ \mathbf{1}_A(x)\ \lambda(dx).$$

Théorème VI.1.10

$f : \mathbb{R} \to \mathbb{R}_+$ borélienne est une densité de probabilité si et seulement si $\int_{\mathbb{R}} f(x)\ \lambda(dx) = 1$.
La mesure de probabilité est alors entièrement déterminée par f.

Variable aléatoire

Définition VI.2.1

Soient $(\Omega, \mathcal{F}, \mathbf{P})$ *un espace de probabilité et* E *un espace muni d'une tribu* $\mathcal{E}$. *On appelle* **variable aléatoire** *à valeurs dans* E *toute application mesurable* $X : \Omega \to E$, *i. e. telle que*

$$\forall A \in \mathcal{E}, \quad X^{-1}(A) \in \mathcal{F}.$$

Variable aléatoire

Définition VI.2.1

Soient $(\Omega, \mathcal{F}, \mathbf{P})$ *un espace de probabilité et* E *un espace muni d'une tribu* $\mathcal{E}$*. On appelle* **variable aléatoire** *à valeurs dans* E *toute application mesurable* $X : \Omega \to E$*, i. e. telle que*

$$\forall A \in \mathcal{E}, \quad X^{-1}(A) \in \mathcal{F}.$$

Définition VI.2.2

Soit $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ *une variable aléatoire. La sous-tribu* $X^{-1}(\mathcal{E}) = \left\{X^{-1}(A);\ A \in \mathcal{E}\right\}$ *de* $\mathcal{F}$ *est appelée* **tribu engendrée par** X*, et est notée* $\sigma(X)$*.*

Loi d'une variable aléatoire

Définition VI.2.3

Soit $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ *une variable aléatoire.*
L'application $P_X : \mathcal{E} \to [0,1]$ *telle que*

$$\forall A \in \mathcal{E}, \quad P_X(A) = \mathbf{P}(\{\omega : X(\omega) \in A\}) = \mathbf{P}(X^{-1}(A))$$

définit une mesure de probabilité sur l'espace $(E, \mathcal{E})$*, appelé* **mesure image** *de* $\mathbf{P}$ *par* X*.*
La mesure de probabilité P_X *est appelée* **distribution** *ou* **loi de** X*.*

On note $\mathbf{P}(X \in A)$ la quantité $P_X(A)$.

Construction de $\boxed{\mathcal{I}_{\mathbf{P}}(X) = \int_{\Omega} X(\omega)\ \mathbf{P}(d\omega)}$ **pour une variable aléatoire réelle** $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

- Pour $X = \sum_{i=1}^{n} \alpha_i \ \mathbf{1}_{A_i}$, où $A_i \in \mathcal{F}$ et $\alpha_i \in \mathbb{R}_+$ pour tout $1 \leq i \leq n$,

$$\mathcal{I}_{\mathbf{P}}(X) = \int_{\Omega} X(\omega)\ \mathbf{P}(d\omega) = \sum_{i=1}^{n} \alpha_i \ \mathbf{P}(A_i).$$

- Pour tout $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable **positive**,

$$\mathcal{I}_{\mathbf{P}}(X) = \sup \left\{ \mathcal{I}_{\mathbf{P}}(\varphi);\ \varphi \text{ étagée telle que } \varphi \leq X \right\}.$$

- Pour tout $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable,

$$X \in L^1(\Omega, \mathcal{F}, \mathbf{P}) \Leftrightarrow \mathcal{I}_{\mathbf{P}}(|X|) < +\infty.$$

On pose $\boxed{\mathcal{I}_{\mathbf{P}}(X) = \mathcal{I}_{\mathbf{P}}(X^+) - \mathcal{I}_{\mathbf{P}}(X^-)}$, où $X^+ = \max(X, 0)$ et $X^- = -\min(X, 0)$.

Définition VI.2.4

Soit X une variable aléatoire réelle telle que $\int_\Omega |X(\omega)|.\mathbf{P}(d\omega) < \infty$ *(*$\Leftrightarrow X \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$*).*
On dit alors que X admet un moment d'ordre 1, et on pose

$$\mathbf{E}[X] = \int_\Omega X(\omega).\mathbf{P}(d\omega) = \int_\Omega X.d\mathbf{P}$$

La quantité $\mathbf{E}[X]$ *est appelée* **espérance (ou moyenne) de** *X.*

Théorème VI.2.5

Soit $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires réelles sur* $(\Omega, \mathcal{F}, \mathbf{P})$.

- **Convergence monotone** : *Si* $X_n \geq 0$ *pour tout* n *et* X_n *croît vers* X *p.s., alors*
$$\lim_{n\to\infty} \mathbf{E}[X_n] = \mathbf{E}[X].$$
- **Lemme de Fatou** : *Si* $X_n \geq 0$ *pour tout* n, *alors*
$$\mathbf{E}\left[\liminf_{n\to\infty} X_n\right] \leq \liminf_{n\to\infty} \mathbf{E}[X_n].$$
- **Convergence dominée** : *Si* $\lim_{n\to\infty} X_n = X$ *p.s. et s'il existe* $Z \in L^1(\Omega, \mathcal{F}, \mathbf{P})$ *telle que* $|X_n| \leq Z$ *pour tout* n, *alors*
$$\lim_{n\to\infty} \mathbf{E}[X_n] = \mathbf{E}[X].$$

Proposition VI.2.6 (Inégalité de Markov)

Soit $X \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$. *Pour tout réel* $a > 0$, *on a*

$$\mathbf{P}(|X| \geq a) \leq \frac{\mathbf{E}[|X|]}{a}.$$

Théorème VI.2.7 (Théorème de transfert)

Soit $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ une variable aléatoire. La loi de X est la mesure de probabilité P_X sur $(E, \mathcal{E})$ caractérisée par

$$\mathbf{E}[h(X)] = \int_E h(x)\; P_X(dx)$$

pour toute application mesurable $h : E \to \mathbb{R}$ bornée, ou telle que $h(X) \in \mathbf{L}^1(\Omega, \mathcal{F}, \mathbf{P})$.

Rem $\quad \forall A \in \mathcal{F}, \quad \mathbf{P}(X \in A) = P_X(A) = \mathbf{E}[\mathbf{1}_{X \in A}]$

- Identification de loi de v.a. **TD Exercice VI.1.2**

Théorème VI.2.8

Soit X une v.a. réelle dont la loi P_X admet une densité f_X, et soit $h : \mathbb{R} \to \mathbb{R}$ mesurable telle que $\int_{\mathbb{R}} |h(x)| f_X(x) dx < \infty$.
Alors, la variable aléatoire $h(X)$ admet un moment d'ordre 1 et

$$\mathbf{E}[h(X)] = \int_{\mathbb{R}} h(x)\ f_X(x)\ \lambda(dx).$$

Définition VI.2.9

On dit qu'une v.a. réelle X sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$ admet un moment d'ordre $n \geq 1$ si $\int_\Omega |X(\omega)|^n \, \mathbf{P}(d\omega) < \infty$.

On note $X \in \mathbf{L}^n(\Omega, \mathcal{F}, \mathbf{P})$.

Définition VI.2.9

On dit qu'une v.a. réelle X sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$ admet un moment d'ordre $n \geq 1$ si $\int_\Omega |X(\omega)|^n \, \mathbf{P}(d\omega) < \infty$.

On note $X \in \mathbf{L}^n(\Omega, \mathcal{F}, \mathbf{P})$.

Proposition VI.2.10

Soient $0 < p < q$.
On a l'inclusion $\mathbf{L}^q(\Omega, \mathcal{F}, \mathbf{P}) \subset \mathbf{L}^p(\Omega, \mathcal{F}, \mathbf{P})$, autrement dit, pour toute v.a. réelle X sur $(\Omega, \mathcal{F}, \mathbf{P})$,

$$\int_\Omega |X|^q d\mathbf{P} < \infty \Rightarrow \int_\Omega |X|^p d\mathbf{P} < \infty.$$

Définition VI.2.11

Si $X \in \mathbf{L}^2(\Omega, \mathcal{F}, \mathbf{P})$*, alors on peut définir la* **variance de** X

$$\mathrm{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right].$$

La quantité $\sqrt{\mathrm{Var}(X)}$ *est appelée* **écart-type de** X*.*

Proposition VI.2.12

- *Pour tous réels a et b,* $\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$.
- *Pour tout réel* $a > 0$, *on a*

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

*C'est l'***inégalité de Bienaymé-Tchebychev***.*

- *X est constante p.s. si et seulement si* $\mathrm{Var}(X) = 0$.

Variable aléatoire constante

X est constante s'il existe $c \in \mathbb{R}$ tel que

$$\mathbf{P}(X = c) = 1.$$

La loi P_X est alors la **distribution de Dirac** δ_c, $\mathbf{E}[X] = c$ et $\mathrm{Var}(X) = 0$.

Variable aléatoire constante

X est constante s'il existe $c \in \mathbb{R}$ tel que

$$\mathbf{P}(X = c) = 1.$$

La loi P_X est alors la **distribution de Dirac** δ_c, $\mathbf{E}[X] = c$ et $\mathrm{Var}(X) = 0$.

Loi uniforme

X suit une **loi uniforme** sur $[a, b] \subset \mathbb{R}$ si sa loi admet la densité de probabilité $f : \mathbb{R} \to \mathbb{R}$ définie par

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b]; \\ 0 & \text{sinon.} \end{cases}$$

Loi exponentielle

X suit une **loi exponentielle** $\mathcal{E}(\lambda)$ $(\lambda > 0)$ si elle admet la densité de probabilité $f : \mathbb{R} \to \mathbb{R}$

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0; \\ 0 & \text{sinon.} \end{cases}$$

On a alors $\mathbf{E}[X] = \frac{1}{\lambda}$ et $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

Loi exponentielle

X suit une **loi exponentielle** $\mathcal{E}(\lambda)$ $(\lambda > 0)$ si elle admet la densité de probabilité $f : \mathbb{R} \to \mathbb{R}$

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0; \\ 0 & \text{sinon.} \end{cases}$$

On a alors $\mathbf{E}[X] = \frac{1}{\lambda}$ et $\mathrm{Var}(X) = \frac{1}{\lambda^2}$.

Loi gamma

X suit une **loi gamma** $\gamma(p, \lambda)$ $(p > 0$ et $\lambda > 0)$ si sa densité est

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} \frac{\lambda}{\Gamma(p)}(\lambda x)^{p-1} e^{-\lambda x} & \text{si } x \geq 0; \\ 0 & \text{sinon.} \end{cases}$$

On a alors $\mathbf{E}[X] = \frac{p}{\lambda}$ et $\mathrm{Var}(X) = \frac{p}{\lambda^2}$.

Loi normale ou gaussienne

X suit une **loi normale** $\mathcal{N}(m, \sigma^2)$ ($(m, \sigma) \in \mathbb{R} \times \mathbb{R}_+$) si elle admet la densité de probabilité $f : \mathbb{R} \to \mathbb{R}$

$$\forall x \in \mathbb{R}, \quad f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right].$$

On a alors $\mathbf{E}[X] = m$ et $\mathrm{Var}(X) = \sigma^2$.

Retour sur les objectifs de la séance

- je maîtrise les notions d'**espace de probabilité**, de **variable aléatoire (def. VI.2.1)**, de **loi (def. VI.2.3)** ;
- je suis capable de calculer la **probabilité d'un événement**, lorsque la mesure de probabilité est donnée ;
- je maîtrise les notions de **fonction de répartition (def. VI.1.5)** et de **densité de probabilité (def. VI.1.9)** ;
- je sais déterminer la loi d'une variable aléatoire ;
- je suis capable de vérifier qu'une variable aléatoire donnée est mesurable par rapport à une sous-tribu ;
- je suis capable de calculer l'**espérance (def. VI.2.4)** et la **variance (def. VI.2.11)** d'une variable aléatoire, lorsqu'elles existent ;
- je maîtrise l'application du **théorème de transfert (Th. VI.2.7)** (calculs, détermination de lois).