

Compléments d'analyse L2 PMRC : intégration et probabilités

Carl Tipler

UNIVERSITÉ DE BREST

E-mail addresses: `carl.tipler@univ-brest.fr`

Table des matières

| | |
|---|----|
| Chapitre I. Intégration | 5 |
| 1. Rappels sur l'intégrale de Riemann | 5 |
| 2. Intégrales généralisées | 10 |
| 3. Convergence dominée, intégrales à paramètres | 13 |
| Chapitre II. Probabilités | 17 |
| 1. Espaces probabilisés | 17 |
| 2. Variables aléatoires discrètes | 21 |
| 3. Variables aléatoires à densité | 24 |
| 4. Convergences de variables aléatoires | 26 |
| Annexe A. Primitives usuelles | 29 |
| Annexe B. Intégrales généralisées types | 31 |
| Annexe C. Variables aléatoires discrètes usuelles | 33 |
| Annexe D. Variables aléatoires à densité usuelles | 35 |
| Annexe. Bibliographie | 37 |

Intégration

Dans ce chapitre on va commencer par rappeler brièvement la construction de l'intégrale de Riemann des fonctions continues par morceaux et quelques unes de ses propriétés. On s'intéressera particulièrement aux méthodes et techniques de calculs. Ensuite, on passera aux intégrales impropres, et on donnera des critères de convergence de ces dernières. Enfin, on verra le théorème de convergence dominée et ces applications.

1. Rappels sur l'intégrale de Riemann

L'intégrale de Riemann a été introduite par Riemann dans les années 1850. On va se restreindre ici à l'intégration des fonctions continues par morceaux, mais la définition peut être étendue à une plus grande classe de fonctions (voir par exemple [RW2]). Dans tout ce chapitre, a et b sont deux réels, avec $a < b$, et $[a, b]$ est donc un segment de \mathbb{R} non réduit à un singleton. On désigne par \mathbb{K} le corps \mathbb{R} ou \mathbb{C} .

1.1. Intégrale des fonctions en escalier. On commence par définir l'intégrale de fonctions particulièrement simples : les fonctions en escalier.

DÉFINITION I.1.1. Une subdivision de $[a, b]$ est une famille $\sigma = (x_0, \dots, x_n)$ de $[a, b]$ telle que

$$a = x_0 < x_1 < \dots < x_n = b.$$

Le réel strictement positif $\min_{0 \leq i \leq n-1} (x_{i+1} - x_i)$ est appelé le pas de σ .

Une subdivision est une famille, mais par abus de langage on s'autorisera à parler de réunion ou d'inclusion de subdivisions de $[a, b]$. Ainsi, $(0, 1, 2, 5)$ est une subdivision de $[0, 5]$ incluse dans $(0, 1, 2, 3, 4, 5)$, et la réunion de $(0, 1, 2, 5)$ et $(0, 2, 4, 5)$ est $(0, 1, 2, 4, 5)$.

DÉFINITION I.1.2. Une fonction $f : [a, b] \rightarrow \mathbb{K}$ est une fonction en escalier si pour une certaine subdivision (x_0, \dots, x_n) de $[a, b]$, dite adaptée à f , on a pour tout $i \in \llbracket 0, n-1 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ est constante.

REMARQUE I.1.3. Soient f, g deux fonctions en escalier sur $[a, b]$.

- (1) Ajouter un nombre fini de points à une subdivision adaptée à f donne une subdivision adaptée à f .
- (2) L'espace $\mathcal{E}([a, b], \mathbb{K})$ des fonctions en escalier à valeurs dans \mathbb{K} est un \mathbb{K} -espace vectoriel stable par produit, par $f \mapsto |f|$ (et par $f \mapsto \Re(f)$ et $f \mapsto \Im(f)$ dans le cas complexe).

DÉFINITION I.1.4. Soient $f : [a, b] \rightarrow \mathbb{K}$ une fonction en escalier, $\sigma = (x_0, \dots, x_n)$ une subdivision de $[a, b]$ adaptée à f , et pour tout $i \in \llbracket 0, n-1 \rrbracket$, y_i la valeur de f sur $]x_i, x_{i+1}[$. Le scalaire

$$\int_a^b f(t) dt := \sum_{i=0}^{n-1} y_i (x_{i+1} - x_i) \in \mathbb{K}$$

ne dépend pas de la subdivision adaptée choisie, et est appelé intégrale de f sur $[a, b]$.

EXEMPLE I.1.5. Calculer $\int_0^{n+1} [t]^2 dt$ pour $n \in \mathbb{N}$ et où $[\cdot]$ désigne la partie entière.

On a les propriétés suivantes :

PROPOSITION I.1.6. *L'intégrale des fonctions en escalier vérifie, pour $(\lambda, \mu) \in \mathbb{K}^2$, $(f, g) \in \mathcal{E}([a, b], \mathbb{K})$:*

(1) *linéarité :*

$$\int_a^b (\lambda f(t) + \mu g(t)) dt = \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt,$$

(2) *inégalité triangulaire :*

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt,$$

(3) *relation de Chasles :*

$$\forall c \in [a, b], \int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt,$$

(4) *dans le cas complexe :*

$$\int_a^b f(t) dt = \int_a^b \Re(f)(t) dt + i \int_a^b \Im(f)(t) dt.$$

Si on suppose de plus que $\mathbb{K} = \mathbb{R}$, alors

(5) *Si $f \geq 0$,*

$$\int_a^b f(t) dt \geq 0,$$

(6) *Si $f \leq g$, alors*

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt.$$

1.2. Intégrale des fonctions continues par morceaux. L'intégrale des fonctions en escalier va être utilisée pour définir celle des fonctions continues par morceaux à l'aide d'approximations.

DÉFINITION I.1.7. Une fonction $f : [a, b] \rightarrow \mathbb{K}$ est une fonction continue par morceaux si pour une certaine subdivision (x_0, \dots, x_n) de $[a, b]$, dite adaptée à f , on a pour tout $i \in \llbracket 0, n-1 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ est continue et admet un prolongement par continuité en x_i et x_{i+1} .

On rappelle que pour une fonction continue f sur $[a, x_0[$, admettre un prolongement par continuité en x_0 signifie que la limite à gauche de la fonction $f(x)$ en x_0 existe (le prolongement $\tilde{f} : [a, x_0] \rightarrow \mathbb{K}$ est alors défini par $\tilde{f}(x_0) = \lim_{x \rightarrow x_0} f(x)$).

REMARQUE I.1.8. L'espace $\mathcal{C}_m([a, b], \mathbb{K})$ des fonctions continues par morceaux à valeurs dans \mathbb{K} est un \mathbb{K} -espace vectoriel stable par produit, par $f \mapsto |f|$ (et par $f \mapsto \Re(f)$ et $f \mapsto \Im(f)$ dans le cas complexe). Il contient le sous espace $\mathcal{E}([a, b], \mathbb{K})$.

On va noter dans la suite, pour toute fonction $f \in \mathcal{C}_m([a, b], \mathbb{K})$, la norme infinie de f par

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)| \in \mathbb{R}_+.$$

On rappelle que par le “théorème des bornes”, la norme infinie est bien définie. De plus, on a l’inégalité triangulaire, pour $f, g \in \mathcal{C}_m([a, b], \mathbb{K})$:

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

La norme infinie définit alors une distance sur l’espace des fonctions continues par morceaux, et la convergence des suites pour cette distance est appelée *convergence uniforme* des fonctions. Un résultat essentiel dans la construction de l’intégrale de Riemann des fonctions continues par morceaux est le suivant :

THÉORÈME I.1.9. *Toute fonction de $\mathcal{C}_m([a, b], \mathbb{K})$ est limite uniforme d’une suite de fonctions de $\mathcal{E}([a, b], \mathbb{K})$.*

Autrement dit, l’espace des fonctions en escalier est “dense” dans celui des fonctions continues par morceaux pour la distance uniforme. Pour démontrer ce théorème, il suffit de se restreindre au cas d’une fonction continue sur un segment, puis d’utiliser le théorème de Heine qui stipule qu’une telle fonction est *uniformément continue*.

REMARQUE I.1.10. Dans le cas réel, une fonction $f \in \mathcal{C}_m([a, b], \mathbb{R})$ peut être approchée par une suite *croissante* de fonctions en escalier *inférieures* à f . Ainsi, dans le cas où f est continue sur $[a, b]$, si on pose pour tout $n \in \mathbb{N}^*$ et tout $i \in \llbracket 0, n \rrbracket$, $x_i := a + i \frac{b-a}{n}$, la suite $(\phi_n)_{n \in \mathbb{N}^*}$ définie par :

$$(1) \quad \phi_n(x) = \sum_{i=0}^{n-1} \left(\inf_{x \in]x_i, x_{i+1}[} f(x) \right) \mathbf{1}_{[x_i, x_{i+1}[}(x)$$

est une suite croissante qui tend uniformément vers f par valeurs inférieures. De même, la suite $(\psi_n)_{n \in \mathbb{N}^*}$ définies par

$$(2) \quad \psi_n(x) = \sum_{i=0}^{n-1} \left(\sup_{x \in]x_i, x_{i+1}[} f(x) \right) \mathbf{1}_{[x_i, x_{i+1}[}(x)$$

est une suite décroissante qui tend uniformément vers f par valeurs supérieures. On remarque par ailleurs que si f est positive, alors pour tout $n \in \mathbb{N}^*$, ϕ_n et ψ_n le sont également. On laisse le soin au lecteur d’adapter ces suites aux cas des fonctions continues par morceaux.

On peut alors définir l’intégrale de Riemann d’une fonction continue par morceaux :

THÉORÈME I.1.11. *Soit $f \in \mathcal{C}_m([a, b], \mathbb{K})$. Pour toute suite $(f_n)_{n \in \mathbb{N}} \in \mathcal{E}([a, b], \mathbb{K})^{\mathbb{N}}$ de fonctions en escalier qui converge uniformément vers f , la limite de la suite $(\int_a^b f_n(t) dt)_{n \in \mathbb{N}}$ existe. Cette limite ne dépend pas de la suite choisie.*

DÉFINITION I.1.12. Soit $f \in \mathcal{C}_m([a, b], \mathbb{K})$. On définit l’intégrale de f sur $[a, b]$ comme étant la limite de $\int_a^b f_n(t) dt$ pour une suite de fonctions en escalier $(f_n)_{n \in \mathbb{N}} \in \mathcal{E}([a, b], \mathbb{K})^{\mathbb{N}}$ qui converge uniformément vers f .

Le théorème précédent implique que l’intégrale est bien définie et coïncide avec la définition de l’intégrale des fonctions en escalier si f est en escalier. Sa preuve repose sur les propriétés de l’intégrale des fonctions en escalier (inégalité triangulaire, croissance et linéarité), ainsi que sur le théorème de Bolzano-Weierstrass pour l’existence de la limite.

Par passages à la limite, on étend les propriétés de l'intégrale aux fonctions continues par morceaux :

PROPOSITION I.1.13. *L'intégrale des fonctions continues par morceaux vérifie, pour $(\lambda, \mu) \in \mathbb{K}^2$, $(f, g) \in \mathcal{C}_m([a, b], \mathbb{K})$:*

(1) *linéarité :*

$$\int_a^b (\lambda f(t) + \mu g(t)) dt = \lambda \int_a^b f(t) dt + \mu \int_a^b g(t) dt,$$

(2) *inégalité triangulaire :*

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt,$$

(3) *relation de Chasles :*

$$\forall c \in [a, b], \int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt,$$

(4) *dans le cas complexe :*

$$\int_a^b f(t) dt = \int_a^b \Re(f)(t) dt + i \int_a^b \Im(f)(t) dt,$$

(5) *si $f = g$ sauf en un nombre fini de points,*

$$\int_a^b f(t) dt = \int_a^b g(t) dt.$$

Si on suppose de plus que $\mathbb{K} = \mathbb{R}$, alors

(6) *si $f \geq 0$,*

$$\int_a^b f(t) dt \geq 0,$$

(7) *si $f \leq g$, alors*

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt,$$

(8) *si $f \geq 0$, et strictement positive en au moins un point de continuité, alors*

$$\int_a^b f(t) dt > 0,$$

(9) *si $f \geq 0$, f est continue, et si $\int_a^b f(t) dt = 0$, alors $f = 0$.*

Parmi les suites de fonctions en escaliers qui convergent uniformément vers f et que l'on peut utiliser pour définir l'intégrale de f , certaines ont une forme particulièrement simple. C'est le cas par exemple des suites $(\phi_n)_{n \in \mathbb{N}^*}$ et $(\psi_n)_{n \in \mathbb{N}^*}$ définies par (1) et (2). D'autre part, avec les notations de la Remarque I.1.10, les fonctions

$$\tau_n^-(x) := \sum_{i=0}^{n-1} f(x_i) \mathbf{1}_{[x_i, x_{i+1}]}(\mathbf{x})$$

et

$$\tau_n^+(x) = \sum_{i=0}^{n-1} f(x_{i+1}) \mathbf{1}_{[x_i, x_{i+1}]}(\mathbf{x})$$

sont des fonctions en escalier qui définissent des suites qui convergent uniformément vers f (elle vérifient $\phi \leq \tau_n^{-,+} \leq \psi$). On en déduit la convergence des *sommes de Riemann* vers l'intégrale de f :

PROPOSITION I.1.14. *Pour toute fonction $f \in \mathcal{C}_m([a, b], \mathbb{K})$,*

$$\int_a^b f(t)dt = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right) = \lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{i=1}^n f\left(a + i \frac{b-a}{n}\right).$$

EXEMPLE I.1.15. Par continuité de $x \mapsto \frac{1}{x+1}$ sur $[0, 1]$, on a

$$\begin{aligned} \sum_{k=1}^n \frac{1}{n+k} &= \frac{1}{n} \sum_{k=1}^n \frac{1}{1 + \frac{k}{n}} \\ &\xrightarrow{n \rightarrow +\infty} \int_0^1 \frac{dt}{1+t}. \end{aligned}$$

On verra dans la section suivante que par le théorème fondamentale de l'analyse cette intégrale vaut $\ln(2)$.

REMARQUE I.1.16. De manière plus générale, on dit qu'une fonction $f : [a, b] \rightarrow \mathbb{K}$ est intégrable au sens de Riemann si

$$\sup\left\{\int_a^b \phi(t)dt, \phi \in \mathcal{E}([a, b], \mathbb{K}), \phi \leq f\right\} = \inf\left\{\int_a^b \psi(t)dt, \psi \in \mathcal{E}([a, b], \mathbb{K}), f \leq \psi\right\}.$$

Dans ce cas, l'intégrale de f est par définition

$$\int_a^b f(t)dt = \sup\left\{\int_a^b \phi(t)dt, \phi \in \mathcal{E}([a, b], \mathbb{K}), \phi \leq f\right\}.$$

Avec la Remarque I.1.10 et la croissance de l'intégrale on déduit que les fonctions continues par morceaux sont intégrables au sens de Riemann et que les deux définitions coïncident. Il est remarquable qu'il existe des fonctions intégrables qui ne sont pas continues par morceaux¹.

1.3. Calculs d'intégrales. On liste ici des propositions utiles pour déterminer la valeur d'une intégrale. On notera I un intervalle de \mathbb{R} , en supposant $a, b \in I^2$, et on notera $\mathcal{C}^k(I, \mathbb{K})$ l'espace des fonctions de classe \mathcal{C}^k sur I et à valeurs dans \mathbb{K} . On pose pour toute fonction $f \in \mathcal{C}_m([a, b], \mathbb{K})$:

$$\int_b^a f(t)dt := - \int_a^b f(t)dt.$$

THÉORÈME I.1.17 (Théorème fondamental de l'analyse). *Soit $f \in \mathcal{C}(I, \mathbb{K})$. La fonction $x \mapsto \int_a^x f(t)dt$ est l'unique primitive de f sur I qui s'annule en a . Pour toute primitive F de f on a*

$$\int_a^b f(t)dt = F(b) - F(a).$$

On utilisera la notation

$$\int_a^b f(t)dt = [F]_a^b = F(b) - F(a).$$

On pourra consulter l'Annexe A à toute fin utile.

1. Les fonctions intégrables au sens de Riemann on un ensemble de points de discontinuité de "mesure nulle".

REMARQUE I.1.18. On déduit de ce théorème que pour toute fonction $f \in \mathcal{C}([a, b], \mathbb{K})$,

$$\int_a^b f(t)dt = \lim_{x \rightarrow b} \int_a^x f(t)dt.$$

C'est ce point de vue qui sera utilisée pour définir les intégrales impropres dans la prochaine section.

À l'aide du théorème fondamental de l'analyse on peut obtenir les corollaires très utiles suivants.

COROLLAIRE I.1.19 (Intégrale, parité et périodicité). *Soit $f \in \mathcal{C}([-a, a], \mathbb{R})$ (on suppose ici $a > 0$).*

(1) *Si f est paire, alors $\int_{-a}^a f(t)dt = 2 \int_0^a f(t)dt$.*

(2) *Si f est impaire, alors $\int_{-a}^a f(t)dt = 0$.*

Soit $f \in \mathcal{C}(\mathbb{R}, \mathbb{R})$. Si f est périodique de période T , alors $\int_a^{a+T} f(t)dt = \int_0^T f(t)dt$.

COROLLAIRE I.1.20 (Intégration par parties). *Soient $u, v \in \mathcal{C}^1(I, \mathbb{K})$. Alors*

$$\int_a^b u'(t)v(t)dt = [uv]_a^b - \int_a^b u(t)v'(t)dt.$$

COROLLAIRE I.1.21 (Changement de variable). *Soient $\phi \in \mathcal{C}^1(I, \mathbb{R})$, $f \in \mathcal{C}(J, \mathbb{C})$ telles que $\phi(I) \subset J$. Alors*

$$\int_a^b f(\phi(t))\phi'(t)dt = \int_{\phi(a)}^{\phi(b)} f(x)dx.$$

2. Intégrales généralisées

On va maintenant s'intéresser aux intégrales "impropres", ou "généralisées". Comme dans la Section 1, on pose $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . On fixe ici a et b deux éléments de la droite complétée $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$, avec $a < b$ (par convention, pour tout $x \in \mathbb{R}$, $-\infty < x < +\infty$). On fera référence à a et b comme étant les *bornes* de l'intervalle $]a, b[$.

2.1. Définition et premières propriétés. On va chercher à calculer l'intégrale de fonctions continues par morceaux sur $]a, b[$.

DÉFINITION I.2.1. Une fonction $f :]a, b[\rightarrow \mathbb{K}$ est continue par morceaux s'il existe une subdivision $\sigma = (x_0 = a, \dots, x_n = b)$ de $]a, b[$ telle que :

- (i) Pour tout $i \in \llbracket 0, n-1 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ est continue,
- (ii) Pour tout $i \in \llbracket 1, n-2 \rrbracket$, la restriction de f à $]x_i, x_{i+1}[$ admet un prolongement par continuité à $[x_i, x_{i+1}]$,
- (iii) La restriction de f à $]a, x_1[$ admet un prolongement par continuité en x_1 ,
- (iv) La restriction de f à $]x_{n-1}, b[$ admet un prolongement par continuité en x_{n-1} .

REMARQUE I.2.2. Si jamais a (resp. b) est un réel, et si $f : [a, b[\rightarrow \mathbb{K}$ (resp. $f :]a, b] \rightarrow \mathbb{K}$), on dira que $f \in \mathcal{C}_m([a, b[, \mathbb{K})$ (resp. $f \in \mathcal{C}_m(]a, b], \mathbb{K})$) si f est continue par morceaux sur $]a, b[$ et la restriction de f à $]a, x_1[$ (resp. $]x_{n-1}, b[$) admet un prolongement par continuité en a (resp. en b). Dans la suite, si on pose $f \in \mathcal{C}_m([a, b[, \mathbb{K})$ (resp. $f \in \mathcal{C}_m(]a, b], \mathbb{K})$, **on sous-entend que $a \in \mathbb{R}$ (resp. $b \in \mathbb{R}$)**. Les espaces $\mathcal{C}_m([a, b[, \mathbb{K})$, $\mathcal{C}_m(]a, b], \mathbb{K})$ et $\mathcal{C}_m(]a, b[, \mathbb{K})$ sont des \mathbb{K} -espaces vectoriels.

DÉFINITION I.2.3. Soit $f \in \mathcal{C}_m([a, b[, \mathbb{K})$ (resp. $f \in \mathcal{C}_m(]a, b], \mathbb{K})$). On dit que l'intégrale impropre (ou généralisée) $\int_a^b f(t)dt$ est *convergente* si la limite

$$\lim_{x \rightarrow b} \int_a^x f(t)dt$$

existe et est finie (resp. $\lim_{x \rightarrow a} \int_x^b f(t)dt$ existe et est finie). On pose dans ce cas

$$\int_a^b f(t)dt := \lim_{x \rightarrow b} \int_a^x f(t)dt$$

(resp. $\int_a^b f(t)dt := \lim_{x \rightarrow a} \int_x^b f(t)dt$). Sinon, cette intégrale est dite *divergente*.

On a aussi la notion suivante :

DÉFINITION I.2.4. Soit $f \in \mathcal{C}_m(]a, b[, \mathbb{K})$. On dit que l'intégrale impropre (ou généralisée) $\int_a^b f(t)dt$ est *convergente* si pour $x_0 \in]a, b[$, les intégrales impropres $\int_a^{x_0} f(t)dt$ et $\int_{x_0}^b f(t)dt$ sont convergentes. On pose dans ce cas

$$\int_a^b f(t)dt := \int_a^{x_0} f(t)dt + \int_{x_0}^b f(t)dt.$$

REMARQUE I.2.5. La définition qui précède ne dépend pas du point x_0 choisi par la relation de Chasles de l'intégrale des fonctions continues par morceaux.

LEMME I.2.6. Soit $f \in \mathcal{C}([a, b[, \mathbb{K})$, de primitive F . Alors $\int_a^b f(t)dt$ est convergente si et seulement si F admet une limite en b , auquel cas

$$\int_a^b f(t)dt = \lim_{x \rightarrow b} F(x) - F(a).$$

Le cas où $f \in \mathcal{C}_m(]a, b], \mathbb{K})$ est symétrique.

EXEMPLES I.2.7. Montrer les faits suivants :

- (1) L'intégrale $\int_0^{+\infty} e^{-t}dt$ est convergente,
- (2) L'intégrale $\int_0^1 \frac{dx}{x}$ est divergente,
- (3) L'intégrale $\int_0^{+\infty} \cos(t)dt$ est divergente,
- (4) L'intégrale $\int_{-\infty}^{+\infty} \frac{dx}{1+x^2}$ est convergente.

La linéarité de la limite et de l'intégrale de Riemann impliquent :

PROPOSITION I.2.8. Les intégrales généralisées vérifient, pour $(\lambda, \mu) \in \mathbb{K}^2$, $(f, g) \in \mathcal{C}_m([a, b[, \mathbb{K})^2$ (resp. $(f, g) \in \mathcal{C}_m(]a, b], \mathbb{K})^2$, $(f, g) \in \mathcal{C}_m(]a, b[, \mathbb{K})^2$) :

- (1) *Linéarité* : si $\int_a^b f(t)dt$ et $\int_a^b g(t)dt$ sont convergentes alors $\int_a^b (\lambda f(t) + \mu g(t))dt$ est convergente et

$$\int_a^b (\lambda f(t) + \mu g(t))dt = \lambda \int_a^b f(t)dt + \mu \int_a^b g(t)dt,$$

- (2) *Inégalité triangulaire* : si $\int_a^b |f(t)|dt$ est convergente, alors $\int_a^b f(t)dt$ est convergente et

$$\left| \int_a^b f(t)dt \right| \leq \int_a^b |f(t)|dt,$$

(3) Relation de Chasles : si $\int_a^b f(t)dt$ est convergente alors

$$\forall c \in]a, b[, \int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt,$$

(4) Dans le cas complexe : $\int_a^b f(t)dt$ est convergente si et seulement si $\int_a^b \Re(f)(t)dt$ et $\int_a^b \Im(f)(t)dt$ le sont, auquel cas

$$\int_a^b f(t)dt = \int_a^b \Re(f)(t)dt + i \int_a^b \Im(f)(t)dt.$$

DÉFINITION I.2.9. Soit $f \in \mathcal{C}_m([a, b[, \mathbb{K})$ (resp. $f \in \mathcal{C}_m(]a, b], \mathbb{K})$, $f \in \mathcal{C}_m(]a, b[, \mathbb{K})$). On dira que $\int_a^b f(t)dt$ est absolument convergente (si $\mathbb{K} = \mathbb{R}$), ou convergente en module (si $\mathbb{K} = \mathbb{C}$) si $\int_a^b |f(t)|dt$ est convergente. On dira dans ce cas que f est *intégrable*.

On a alors trois possibilités :

- (1) Soit f est intégrable,
- (2) ou $\int_a^b f(t)dt$ converge mais $\int_a^b |f(t)|dt$ diverge : on dit que l'intégrale est *semi-convergente*,
- (3) ou $\int_a^b f(t)dt$ est divergente.

2.2. Critères de convergence : comparaisons. On va établir, comme pour les séries, des résultats de convergence par comparaisons. Pour utiliser ces résultats, on retiendra les exemples classiques en Annexe B.

PROPOSITION I.2.10. Soit $f \in \mathcal{C}_m([a, b[, \mathbb{R})$. On suppose qu'il existe $x_0 \in [a, b[$ tel que pour tout $x \in [x_0, b[$, $f(x) \geq 0$. Alors soit $\int_a^b f(t)dt$ est convergente, soit elle est divergente vers $+\infty$.

COROLLAIRE I.2.11. Soient $f, g \in \mathcal{C}_m([a, b[, \mathbb{R})$. On suppose qu'il existe $x_0 \in [a, b[$ tel que pour tout $x \in [x_0, b[$, $g(x) \geq f(x) \geq 0$. Alors :

- (1) Si $\int_a^b g(t)dt$ est convergente, $\int_a^b f(t)dt$ est convergente,
- (2) Si $\int_a^b f(t)dt$ est divergente, $\int_a^b g(t)dt$ est divergente.

Un corollaire du corollaire qui est extrêmement utile :

COROLLAIRE I.2.12. Soient $f, g \in \mathcal{C}_m([a, b[, \mathbb{R})$. On suppose qu'il existe $x_0 \in [a, b[$ tel que pour tout $x \in [x_0, b[$, $g(x) \geq 0$ et $f(x) \geq 0$. Alors :

- (1) Si $f(x) \underset{x \rightarrow b}{\sim} g(x)$, alors les intégrales $\int_a^b g(t)dt$ et $\int_a^b f(t)dt$ sont de même natures,
- (2) Si $f(x) \underset{x \rightarrow b}{=} \mathcal{O}(g(x))$ et si $\int_a^b g(t)dt$ est convergente, $\int_a^b f(t)dt$ est convergente,
- (3) Si $f(x) \underset{x \rightarrow b}{=} \mathcal{O}(g(x))$ et si $\int_a^b f(t)dt$ est divergente, $\int_a^b g(t)dt$ est divergente.

DÉMONSTRATION. Il suffit d'appliquer le Corollaire I.2.11, en notant que les équivalences et \mathcal{O} impliquent des encadrements. Ainsi, $f(x) \underset{x \rightarrow b}{\sim} g(x)$ implique qu'au voisinage de b on a $\frac{1}{2}f \leq g \leq 2f$. □

REMARQUE I.2.13. On laissera au lecteur le soin de formuler des résultats analogues pour des fonctions $f, g \in \mathcal{C}_m(]a, b], \mathbb{R})$, ou pour des fonction négatives à partir d'un réel $x_0 \in]a, b[$.

On termine cette section par le fameux critère de comparaison “séries-intégrales”. Ce dernier permet de déterminer la nature de certaines séries numériques. On suppose ici $a \geq 0$ et f à **valeurs positives**.

THÉORÈME I.2.14. *Soit $f \in \mathcal{C}_m([a, +\infty[, \mathbb{R}_+)$, avec $a \geq 0$. On suppose que f est monotone. Alors l'intégrale $\int_a^{+\infty} f(t)dt$ et la série $\sum_{n \geq a} f(n)$ sont de même nature.*

DÉMONSTRATION. Si par exemple f est décroissante, on a pour tout entier $k \geq [a] + 1$:

$$f(k) \geq \int_k^{k+1} f(t)dt \geq f(k+1).$$

On peut alors sommer ces termes, et utiliser les comparaisons de suites positives. □

3. Convergence dominée, intégrales à paramètres

De nombreux procédés d'analyse font apparaître des limites de fonctions (approximations, séries de Fourier, séries entières, etc) . Une question naturelle est alors de savoir, pour une suite de fonctions $(f_n)_{n \in \mathbb{N}}$ définies sur un intervalle $I \subset \mathbb{R}$ qui converge simplement vers f , si $\int_I f_n(t)dt$ converge vers $\int_I f(t)dt$, ou encore si on peut “intervertir limite et intégrale”. Sans hypothèse supplémentaire la réponse est non en général. Le *théorème de convergence dominée* permet d'intervertir la limite et l'intégrale sous des hypothèses très souples. Ce théorème fondamental a de très nombreux corollaires, eux mêmes utiles dans l'étude des “intégrales à paramètres” ou l'étude des séries de fonctions. Dans tout cette section, I désigne un intervalle de \mathbb{R} et $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

3.1. Convergence dominée et premières conséquences.

THÉORÈME I.3.1 (Théorème de convergence dominée). *Soit $(f_n)_{n \in \mathbb{N}} \in \mathcal{C}_m(I, \mathbb{K})^{\mathbb{N}}$ une suite de fonctions continues par morceaux. On suppose que :*

- (i) *La suite $(f_n)_{n \in \mathbb{N}}$ converge simplement vers une fonction $f \in \mathcal{C}_m(I, \mathbb{K})$,*
- (ii) *Il existe une fonction **intégrable** $\phi \in \mathcal{C}_m(I, \mathbb{K})$ telle que pour tout $n \in \mathbb{N}$ et tout $x \in I$,*

$$|f_n(x)| \leq \phi(x).$$

Alors f est intégrable et

$$\lim_{n \rightarrow +\infty} \int_I f_n(t)dt = \int_I f(t)dt.$$

Notez que l'hypothèse de domination implique que pour tout $n \in \mathbb{N}$, f_n est intégrable. Le théorème stipule alors que

$$\lim_{n \rightarrow +\infty} \int_I f_n(t)dt = \int_I \left(\lim_{n \rightarrow \infty} f_n(t) \right) dt.$$

REMARQUE I.3.2. Pour appliquer ce résultat il ne faudra pas oublier de vérifier que toutes les fonctions qui interviennent sont continues par morceaux. C'est en effet nécessaire dans notre cadre d'intégration, simplement pour donner un sens aux intégrales.

Un cadre particulièrement simple d'application de ce résultat est le suivant :

COROLLAIRE I.3.3. Soit $(f_n)_{n \in \mathbb{N}} \in \mathcal{C}([a, b], \mathbb{K})$ une suite de fonctions **continues** qui converge **uniformément** vers f . Alors

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt.$$

On a aussi, pour les séries, le corollaire suivant ² :

COROLLAIRE I.3.4. *Théorème d'interversion série-intégrale* Soit $(f_n)_{n \in \mathbb{N}} \in \mathcal{C}_m(I, \mathbb{K})$. On suppose que :

(i) La série $\sum_{n=0}^{+\infty} f_n$ converge simplement vers une fonction continue par morceaux sur I ,

(ii) La série numérique $\sum_{n=0}^{+\infty} \int_I |f_n(t)| dt$ est convergente.

Alors $\sum_{n=0}^{+\infty} f_n$ est intégrable et

$$\int_I \left(\sum_{n=0}^{+\infty} f_n \right) dt = \sum_{n=0}^{+\infty} \left(\int_I f_n(t) dt \right).$$

Dans les hypothèses de ce corollaire, l'hypothèse (i) correspond à l'hypothèse (i) du théorème de convergence dominée, et permet de parler des intégrales considérées. L'hypothèse (ii) implique ³ l'hypothèse de domination du théorème de convergence dominée. Elle implique en particulier que pour tout $n \in \mathbb{N}$, f_n est intégrable sur I .

REMARQUE I.3.5. Ce théorème d'interversion "terme à terme" est également une conséquence du théorème de Fubini, dont on verra une version édulcorée dans la section suivante.

3.2. Intégrales à paramètres. Dans cette section, on va s'intéresser à des fonctions définies par des intégrales, qu'on appelle encore intégrales à paramètres. Soient $A \subset E$ une partie d'un espace vectoriel normé $(E, \|\cdot\|)$ (e.g. E un intervalle de \mathbb{R}) et $I \subset \mathbb{R}$ un intervalle de \mathbb{R} . On pose

$$\begin{aligned} f : A \times I &\rightarrow \mathbb{K} \\ (x, t) &\mapsto f(x, t) \end{aligned}$$

une fonction à valeurs dans \mathbb{K} et on souhaiterait définir

$$\begin{aligned} F : A &\rightarrow \mathbb{K} \\ x &\mapsto \int_I f(x, t) dt. \end{aligned}$$

Bien entendu, si pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux et intégrable, alors F est bien définie. On va présenter deux résultats de régularité pour de telles fonctions F . Ces théorèmes reposent sur le théorème de convergence dominée. On les énonce avec les notations qui précèdent.

2. Ce n'est pas un simple corollaire du théorème de convergence dominée. Sa preuve standard repose également sur le théorème de convergence monotone de Lebesgue que nous n'aborderons pas dans ce cours.

3. C'est ici qu'il faut un peu de travail, i.e. le théorème de convergence monotone.

THÉORÈME I.3.6 (Théorème de continuité des intégrales à paramètres). *On suppose que*

- (i) *Pour tout $x \in A$, $t \mapsto f(x, t)$ est continue par morceaux sur I ,*
- (ii) *Pour tout $t \in I$, $x \mapsto f(x, t)$ est continue sur A ,*
- (iii) *Il existe une fonction $\phi \in \mathcal{C}_m(I, \mathbb{R}_+)$ **intégrable** sur I telle que*

$$\forall (x, t) \in A \times I, |f(x, t)| \leq \phi(t).$$

Alors F est bien définie et continue sur A .

L'hypothèse de domination (hypothèse (iii)) implique que $t \mapsto f(x, t)$ est intégrable pour tout $x \in A$ et donc que F est bien définie. Le contenu du théorème est la continuité de F , qui s'obtient par convergence dominée en utilisant la caractérisation séquentielle de la continuité. En pratique, pour obtenir la domination, on préfère se restreindre aux parties compactes de A (un segment $[a, b]$ si A est un intervalle de \mathbb{R}), et montrer ainsi la continuité de F sur tout compact, ce qui implique le même résultat. On a aussi le cas particulier suivant :

COROLLAIRE I.3.7. *Si I est un segment et f est continue sur $A \times I$, alors F est continue sur A .*

Pour la dérivabilité de F , on utilisera le résultat suivant, où $A = J \subset \mathbb{R}$ est cette fois-ci un intervalle de \mathbb{R} :

THÉORÈME I.3.8 (Théorème de régularité \mathcal{C}^1 des intégrales à paramètres). *On suppose que :*

- (i) *Pour tout $x \in J$, $t \mapsto f(x, t)$ est continue par morceaux et intégrable sur I ,*
- (ii) *Pour tout $t \in I$, $x \mapsto f(x, t)$ est de classe \mathcal{C}^1 sur J ,*
- (iii) *Pour tout $x \in J$, $t \mapsto \frac{\partial f}{\partial x}(x, t)$ est continue par morceaux sur I ,*
- (iv) *Il existe une fonction $\phi \in \mathcal{C}_m(I, \mathbb{R}_+)$ **intégrable** sur I telle que*

$$\forall (x, t) \in J \times I, \left| \frac{\partial f}{\partial x}(x, t) \right| \leq \phi(t).$$

Alors F est bien définie, de classe \mathcal{C}^1 sur J et pour tout $x \in J$,

$$\frac{\partial F}{\partial x}(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt.$$

Notez ici que l'hypothèse de domination n'est pas suffisante pour montrer que F est bien définie, il faut donc ajouter l'intégrabilité de f par rapport à $t \in I$ en (i). Comme pour le théorème de continuité, pour montrer la domination on pourra se restreindre aux compactes de J . On a également le corollaire suivant sur les segments :

COROLLAIRE I.3.9. *Si I est un segment et f est de classe \mathcal{C}^1 sur $J \times I$, alors F est de classe \mathcal{C}^1 sur J et pour tout $x \in J$,*

$$\frac{\partial F}{\partial x}(x) = \int_I \frac{\partial f}{\partial x}(x, t) dt.$$

Une autre question naturelle se pose quand on étudie les fonctions définies par une intégrale. Si on sait les dériver sous certaines hypothèses, peut-on aussi les intégrer ? Dans cette direction on va se restreindre au cas compact, i.e. I et J deux segments de \mathbb{R} .

THÉORÈME I.3.10 (Théorème de Fubini (compact)). *Supposons que $f : [a, b] \times [c, d] \rightarrow \mathbb{K}$ soit continue. Alors*

$$F(x, t) = \int_c^d f(x, t) dt$$

est continue et intégrable sur $[a, b]$ et

$$\int_a^b F(x, t) dx = \int_c^d \left(\int_a^b f(x, t) dx \right) dt.$$

Le Corollaire I.3.7 implique immédiatement : que F est continue, et donc intégrable car $[a, b]$ est compacte. De même, la fonction $G : t \mapsto \int_a^b f(x, t) dx$ est intégrable sur $[c, d]$. Le contenu du théorème nous dit que les intégrales de F et de G coïncident, c'est à dire qu'on peut intervertir l'ordre d'intégration :

$$\int_a^b \left(\int_c^d f(x, t) dt \right) dx = \int_c^d \left(\int_a^b f(x, t) dx \right) dt.$$

Probabilités

Ce chapitre résume les concepts basiques de la théorie des probabilités. On commencera par le cadre “discret”, dans lequel on introduira les notions d’espace probabilisé, d’indépendance, de variable aléatoire et de moments. On présentera ensuite le cas des variables aléatoires à densité continue par morceaux. Enfin, on donnera les énoncés de la loi faible des grands nombres et du théorème central limite. On renvoie le lecteur intéressé à [C] pour plus de détails.

1. Espaces probabilisés

La théorie moderne des probabilités est basée sur l’approche axiomatique de Kolmogorov¹, elle-même basée sur la théorie de la mesure de Borel² et Lebesgue³. Dans ce chapitre, on se limitera au cadre discret. Afin de modéliser les phénomènes aléatoires, on va introduire la notion d’*espace probabilisé*. Un tel objet est construit à partir de trois ingrédients : *l’univers*, *les évènements*, et *la mesure de probabilité*. L’univers, noté Ω , est l’ensemble de tous les résultats possibles de l’expérience aléatoire que l’on cherche à modéliser.

EXEMPLES II.1.1. Exemples d’univers (encore appelés espaces des observables, ou espaces échantillons) :

- (1) Tirage à pile ou face : $\Omega = \{P, F\}$;
- (2) Suite de lancer de dés : $\Omega = \{(x_n)_{n \in \mathbb{N}} \in \{1, \dots, 6\}^{\mathbb{N}}\}$,
- (3) Deux lancers de dés successifs : $\Omega = \{(m, n) \in \{1, 2, 3, 4, 5, 6\}^2\}$.

Un *évènement* est une propriété dont on peut dire qu’elle est vérifiée ou non une fois le résultat de l’expérience aléatoire connu.

EXEMPLES II.1.2. Exemples d’évènements pour l’expérience aléatoire “On lance deux fois un dé”, avec $\Omega = \{(m, n) \in \{1, 2, 3, 4, 5, 6\}^2\}$:

- (1) “Le second lancé est un 6” :

$$\{(m, 6), m \in \{1, 2, 3, 4, 5, 6\}\};$$

- (2) “Le premier lancer est supérieur au second” :

$$\{(m, n) \in \Omega, m > n\};$$

- (3) “La somme des deux lancers est paire” :

$$\{(m, n) \in \Omega, m + n \equiv 0 [2]\}.$$

1. Andreï Nikolaïevich Kolomogorov (1903 - 1987), Mathématicien Russe

2. Félix Édouard Justin Émil Borel (1871-1956), mathématicien et homme politique français

3. Henri Léon Lebesgue (1875-1941), mathématicien français

L'ensemble Ω est lui-même un évènement, appelé *évènement certain*, tandis que l'ensemble vide \emptyset est appelé *évènement impossible*. Si A est un évènement, on appelle A^c *évènement contraire* de A .

Enfin, la *mesure de probabilité*, ou encore probabilité, est une fonction qui associe à tout évènement une valeur positive qui représente le degré de confiance que l'on a en sa réalisation. Les valeurs possibles seront comprises entre 0 et 1, avec l'interprétation que plus la probabilité est proche de 1, et plus notre confiance en sa réalisation est grande. Un évènement de probabilité 1 sera appelé *évènement presque-sûr*.

REMARQUE II.1.3. Attention cependant, il existe des évènements presque sûrs différents de l'évènement certain, et des évènements de probabilité nulle qui ne sont pas impossibles.

Une manière heuristique, quoique limitée, d'associer à des évènements une probabilité consiste à réaliser une infinité de fois et indépendamment l'expérience, puis de définir la probabilité de l'évènement A comme étant sa fréquence de réalisation asymptotique. Autrement dit, si on réalise n -fois l'expérience, on note $f_n(A)$ le nombre de fois où A est réalisé divisé par n , et on définit la probabilité de A , notée $\mathbb{P}(A)$:

$$\mathbb{P}(A) = \lim_{n \rightarrow +\infty} f_n(A).$$

On observe alors que par définition :

- (i) Pour tout évènement A , $0 \leq \mathbb{P}(A) \leq 1$,
- (ii) $\mathbb{P}(\Omega) = 1$,
- (iii) Pour deux évènements A et B *incompatibles* (i.e. $A \cap B = \emptyset$), on a

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Ce sont ces propriétés qui vont servir à définir de manière axiomatique un espace probabilisé.

REMARQUE II.1.4. On verra qu'il sera souhaitable, et naturel, d'imposer une condition plus forte que (iii) qui sera remplacée par : pour toute suite d'évènements deux à deux incompatibles $(A_n)_{n \in \mathbb{N}}$,

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

1.1. Définitions et premières propriétés. On souhaiterait définir un espace probabilisé comme étant un univers Ω et une probabilité \mathbb{P} définie sur $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω qui représenterait l'ensemble des évènements possibles. Cette approche fonctionne si Ω est fini, mais pose problème si Ω est infini.

REMARQUE II.1.5. Si $\Omega = \{0, 1\}^{\mathbb{N}}$, il n'existe pas d'application $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$ possédant les propriétés suivantes :

- (i) $\mathbb{P}(\Omega) = 1$,
- (ii) Pour toute suite d'évènements deux à deux incompatibles $(A_n)_{n \in \mathbb{N}}$,

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n),$$

- (iii) Pour tout $A \in \mathcal{P}(\Omega)$ et $n \geq 0$, $\mathbb{P}(T_n(A)) = \mathbb{P}(A)$, où l'application $T_n : \Omega \rightarrow \Omega$ est définie par

$$T_n(\omega_0, \omega_1, \dots) = (\omega_0, \dots, 1 - \omega_n, \dots).$$

Cette expérience représente une suite infinie de lancers d'une pièce de monnaie, et le point (iii) assure l'indépendance des lancers, ainsi que le fait que la pièce soit équilibrée. La preuve de ce fait repose sur l'axiome du choix.

On va alors se restreindre pour les événements à une sous-partie de $\mathcal{P}(\Omega)$ qui satisfait des axiomes naturels :

DÉFINITION II.1.6. Soit Ω un univers. On appelle *tribu* sur Ω une famille \mathcal{T} de parties de Ω qui vérifie :

- (i) \mathcal{T} est non vide,
- (ii) Si $A \in \mathcal{T}$, alors $A^c \in \mathcal{T}$,
- (iii) Si $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$, alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{T}$.

On appelle alors *événement* tout élément de \mathcal{T} .

Ces axiomes nous permettent alors de parler d'événements contraires et de réunion d'événements. Le troisième axiome est appelé *σ -additivité*, et une tribu est aussi appelée *σ -algèbre*.

PROPOSITION II.1.7. Soit \mathcal{T} une tribu sur un ensemble Ω . Alors :

- (1) $\Omega \in \mathcal{T}$ et $\emptyset \in \mathcal{T}$,
- (2) Si $(A_n)_{n \in \mathbb{N}} \in \mathcal{T}^{\mathbb{N}}$, alors $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{T}$,
- (3) Si $A, B \in \mathcal{T}$, $A \setminus B \in \mathcal{T}$.

On peut désormais définir un *espace probabilisé* :

DÉFINITION II.1.8. On appelle espace probabilisable un couple (Ω, \mathcal{T}) où Ω est un ensemble non vide et \mathcal{T} est une tribu sur Ω . On appelle espace probabilisé un triplet $(\Omega, \mathcal{T}, \mathbb{P})$ où (Ω, \mathcal{T}) est un espace probabilisable et \mathbb{P} est une probabilité sur (Ω, \mathcal{T}) , c'est à dire une fonction

$$\mathbb{P} : \mathcal{T} \rightarrow [0, 1]$$

qui vérifie :

- (i) $\mathbb{P}(\omega) = 1$,
- (ii) Pour toute suite d'événements deux à deux incompatibles $(A_n)_{n \in \mathbb{N}}$,

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

On peut démontrer :

PROPOSITION II.1.9. Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Alors

- (1) $\mathbb{P}(\emptyset) = 0$,
- (2) Pour tout $A \in \mathcal{T}$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$,
- (3) *Monotonie* : pour tout $A, B \in \mathcal{T}$, si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$,
- (4) *Additivité finie* : si A_1, \dots, A_n sont des événements deux à deux incompatibles,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i),$$

(5) *Sous-additivité* : si I est au plus dénombrable, et si $(A_i)_{i \in I}$ est une collection d'évènements, alors

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i),$$

(6) *Pour tout* $A, B \in \mathcal{T}$,

$$\mathbb{P}(A \cup B) + \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B),$$

(7) *Continuité* : si $(A_n)_{n \in \mathbb{N}}$ est une suite croissante (respectivement décroissante) d'évènements, alors

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n)$$

(respectivement

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n).)$$

Dans toute la suite, $(\Omega, \mathcal{T}, \mathbb{P})$ désigne un espace probablisé. On parlera d'espace probablisé "discret" si Ω est au plus dénombrable.

1.2. Probabilités conditionnelles. Il est fréquent de rencontrer des énoncés du type "si B a lieu, alors la probabilité de A est p ". On parle dans ce cadre de probabilité conditionnelle, où la réalisation de l'évènement B conditionne celle de l'évènement A . En termes de fréquences, si on répète une expérience n fois, on s'intéresse au nombre de fois que A et B sont réalisés, divisé par le nombre de fois où B est réalisé, ce qui vaut $f_n(A \cap B)/f_n(B)$. Un passage à la limite moitive alors :

DÉFINITION II.1.10. Soit $B \in \mathcal{T}$ un évènement tel que $\mathbb{P}(B) > 0$. Pour tout $A \in \mathcal{T}$, la probabilité conditionnelle de A sachant B est la quantité

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

PROPOSITION II.1.11. Soit $B \in \mathcal{T}$ un évènement tel que $\mathbb{P}(B) > 0$. L'application

$$\begin{aligned} \mathbb{P}_B : \mathcal{T} &\rightarrow \mathbb{R} \\ A &\mapsto \mathbb{P}(A|B) \end{aligned}$$

est une mesure de probabilité sur (Ω, \mathcal{T}) .

Les deux résultats suivants, malgré leur simplicité, sont très utiles :

THÉORÈME II.1.12 (Formule de Bayes). Soient $(B_i)_{i \in I} \in \mathcal{T}^I$ une partition au plus dénombrable de Ω , et $A \in \mathcal{T}$. On suppose que pour tout $i \in I$, $\mathbb{P}(B_i) > 0$. Alors

(1) *Loi de probabilité totale* :

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i),$$

(2) *si de plus* $\mathbb{P}(A) > 0$, *la formule de Bayes donne* :

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j \in I} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}.$$

Un cas particulier très utile est le suivant :

COROLLAIRE II.1.13. Soit $(A, B) \in \mathcal{T}$ deux évènements tels que $\mathbb{P}(B) > 0$ et $\mathbb{P}(B^c) > 0$. Alors

$$\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c).$$

Le fait de savoir qu'un évènement B est réalisé à une impacte sur la confiance que l'on a dans la réalisation d'autres évènements. La probabilité d'un autre évènement A est remplacée par la probabilité conditionnelle de A sachant B . Si jamais $\mathbb{P}(A|B) = \mathbb{P}(A)$, le fait de savoir B réalisé n'affecte pas la confiance en la réalisation de A et les évènements sont "indépendants". La définition suivante est valable même si B est de probabilité nulle :

DÉFINITION II.1.14. Deux évènements A et B sont dits indépendants (sous \mathbb{P}) si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Une famille au plus dénombrable d'évènements $(A_i)_{i \in I}$ est indépendante si pour tout sous-ensemble fini $J \subset I$ on a

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

2. Variables aléatoires discrètes

On peut être intéressé par une valeur numérique associée au résultat d'une expérience aléatoire, plus que par le résultat de l'expérience lui-même. Si l'espace probabilisé sous-jacent est discret, on parlera de *variable aléatoire discrète*.

2.1. Définitions. Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé discret.

DÉFINITION II.2.1. Une variable aléatoire discrète est une application $X : \Omega \rightarrow \mathbb{R}$ qui est "mesurable", i.e. qui vérifie :

$$\forall a \in \mathbb{R}, X^{-1}(]-\infty, a]) \in \mathcal{T}.$$

REMARQUE II.2.2. En pratique, dans nos exemples d'espaces probabilisés discrets, on aura $\mathcal{T} = \mathcal{P}(\Omega)$, et donc la mesurabilité sera une condition automatiquement vérifiée. La condition de mesurabilité est nécessaire en général pour donner sens à la définition de fonction de répartition.

Du point de vue probabiliste, on est intéressé par la loi de X :

DÉFINITION II.2.3. Soit $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire discrète. On appelle *loi de X* la mesure de probabilité sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ définie par⁴

$$\forall A \in \mathcal{P}(\mathbb{R}), \mathbb{P}_X(A) = \mathbb{P}(X \in A).$$

La *fonction de masse* de X est la fonction $f_X : \mathbb{R} \rightarrow [0, 1]$ donnée par

$$f_X(x) = \mathbb{P}(X = x).$$

Enfin, la *fonction de répartition* de X est la fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ donnée par

$$F_X(x) = \mathbb{P}(X \leq x).$$

4. Cette définition ne fonctionne que si X est discrète. Pour le cas général, il faudra se restreindre à l'espace probabilisable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, où $\mathcal{B}(\mathbb{R})$ est la tribu des boréliens de \mathbb{R} , i.e. la plus petite tribu de \mathbb{R} qui contient tous les ensembles de la forme $]-\infty, a]$.

La donnée de \mathbb{P}_X permet de déterminer f_X et F_X . Réciproquement, la fonction de masse ou la fonction de répartition suffit à déterminer la loi de X . On a aussi :

PROPOSITION II.2.4. *Soit $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire discrète de fonction de masse f_X et de fonction de répartition F_X . Alors :*

(1) *Pour tout $A \subset \mathbb{R}$ on a*

$$\mathbb{P}_X(A) = \sum_{x \in A \cap X(\Omega)} f_X(x),$$

(2) *F_X est croissante,*

(3) $\lim_{x \rightarrow -\infty} F_X(x) = 0,$

(4) $\lim_{x \rightarrow +\infty} F_X(x) = 1,$

(5) *F_X est continue à droite, i.e. pour tout $x \in \mathbb{R}$, et toute suite décroissante $(x_n)_{n \in \mathbb{N}}$ qui tend par valeurs supérieures vers x , on a*

$$\lim_{n \rightarrow +\infty} F_X(x_n) = F_X(x).$$

On donne des exemples de variables aléatoires discrètes, ainsi que certaines de leurs propriétés, dans l'Annexe C.

DÉFINITION II.2.5. Deux variables aléatoires discrètes X et Y sur $(\Omega, \mathcal{T}, \mathbb{P})$ sont *indépendantes* si pour tout $A, B \subset \mathbb{R}^2$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants. Une famille au plus dénombrable $(X_i)_{i \in I}$ de variables aléatoires discrètes est indépendante si les événements $\{X_j \in A_j\}_{j \in J}$ sont indépendants pour tout $A_j \subset \mathbb{R}$, $j \in J$ et $J \subset I$ fini.

2.2. Moments. À une variable aléatoire on peut associer deux quantités qui caractérisent son comportement : l'espérance et la variance. On pose $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire discrète.

DÉFINITION II.2.6. Supposons que

$$\sum_{x \in X(\Omega)} |x| f_X(x) < +\infty.$$

On définit l'*espérance* de X comme étant

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x f_X(x) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\omega).$$

LEMME II.2.7 (Linéarité de l'espérance). *Soient X et Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{T}, \mathbb{P})$. On suppose que X et Y possèdent une espérance. Alors pour tout $(\alpha, \beta) \in \mathbb{R}^2$, $\alpha X + \beta Y$ possède une espérance, et*

$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y).$$

On a aussi le contrôle suivant :

PROPOSITION II.2.8 (Inégalité de Jensen). *Si X admet une espérance, et si $\phi : \mathbb{R} \rightarrow \mathbb{R}$ est une fonction convexe, alors*

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

En particulier,

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|).$$

Une généralisation de l'espérance est la suivante :

DÉFINITION II.2.9. On appelle $\mathbb{E}(X^n)$ le *moment d'ordre n* , pourvu que cette quantité soit bien définie.

Si l'espérance donne une idée de la moyenne des valeurs que prend X quand on répète une expérience aléatoire, sa variance permet de quantifier à quel point les valeurs de X s'écartent de cette moyenne.

DÉFINITION II.2.10. Supposons que X admette une espérance. On définit la *variance* de X comme étant la quantité

$$\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2).$$

On appelle écart-type de X la quantité

$$\sigma(X) = \sqrt{\text{Var}(X)}.$$

Notez que la variance et l'écart-type peuvent être infinis.

PROPOSITION II.2.11. *La variance de variables aléatoires X et Y vérifient :*

- (1) $\text{Var}(X) \geq 0$ et $\text{Var}(X) = 0$ si et seulement si $\mathbb{P}(X = \mathbb{E}(X)) = 1$,
- (2) Si $\text{Var}(X) < +\infty$, alors $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$,
- (3) Pour $a, b \in \mathbb{R}$, $\text{Var}(a + bX) = b^2\text{Var}(X)$,
- (4) Si $\text{Var}(X) < +\infty$ et $\text{Var}(Y) < +\infty$, alors $\text{Var}(X + Y) < +\infty$.

L'inégalité suivante précise comment la variance contrôle les fluctuations d'une variable aléatoire autour de sa moyenne :

PROPOSITION II.2.12 (Inégalité de Bienaymé-Tchebychev). *Supposons que X admette une espérance. Alors pour tout $a > 0$,*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\text{Var}(X)}{a^2}.$$

La variance, contrairement à l'espérance, n'est pas linéaire.

DÉFINITION II.2.13. On appelle *covariance* de deux variables aléatoires X et Y la quantité

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

On peut vérifier que

$$\text{cov}(X, Y) = \mathbb{E}(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$$

et donc

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{cov}(X, Y).$$

DÉFINITION II.2.14. Deux variables aléatoires X et Y sont *non-corrélées* si $\text{cov}(X, Y) = 0$. Dans ce cas on a

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

LEMME II.2.15. *La covariance est bilinéaire et symétrique en ses variables.*

L'indépendance implique la non-corrélation, la réciproque étant fausse.

PROPOSITION II.2.16. *Deux variables aléatoires indépendantes dont l'espérance existe sont non-corrélées.*

3. Variables aléatoires à densité

De nombreuses expériences aléatoires font intervenir des espaces et des variables non discrètes, par exemple la durée de vie d'une ampoule ($\Omega = \mathbb{R}_+$) l'évolution du cours d'une action sur une période de temps $[s, t]$ ($\Omega = \mathcal{C}^0([s, t], \mathbb{R}_+)$), la trajectoire d'un grain de pollen en suspension ($\Omega = \mathcal{C}^0(\mathbb{R}_+, \mathbb{R}^3)$), etc. Pour étudier de manière axiomatique ce genre d'expérience, il est nécessaire d'autoriser des variables aléatoires définies sur un ensemble non-dénombrable. On va étudier un cas particulier fréquent en pratique, celui de variable aléatoire "à densité". On va voir que les définitions et résultats dans ce cadre sont similaires au cas discret, il suffit essentiellement de remplacer les sommes par des intégrales. Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé.

3.1. Définition.

DÉFINITION II.3.1. Une variable aléatoire est une application $X : \Omega \rightarrow \mathbb{R}$ qui est mesurable, i.e. qui vérifie :

$$\forall a \in \mathbb{R}, X^{-1}(] - \infty, a]) \in \mathcal{T}.$$

On appelle *loi de X* la mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie par ⁵

$$\forall [a, b] \subset \mathbb{R}, \mathbb{P}_X([a, b]) = \mathbb{P}(X \in [a, b]).$$

La *fonction de répartition* de X est la fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ donnée par

$$F_X(x) = \mathbb{P}(X \leq x).$$

La donnée de \mathbb{P}_X permet de déterminer F_X et réciproquement. On a :

PROPOSITION II.3.2. Soit $X : \Omega \rightarrow \mathbb{R}$ une variable aléatoire de fonction de répartition F_X . Alors :

- (1) F_X est croissante,
- (2) $\lim_{x \rightarrow -\infty} F_X(x) = 0$,
- (3) $\lim_{x \rightarrow +\infty} F_X(x) = 1$,
- (4) F_X est continue à droite, i.e. pour tout $x \in \mathbb{R}$, et toute suite décroissante $(x_n)_{n \in \mathbb{N}}$ qui tend par valeurs supérieures vers x , on a

$$\lim_{n \rightarrow +\infty} F_X(x_n) = F_X(x).$$

Réciproquement, toute fonction $F : \Omega \rightarrow \mathbb{R}$ qui vérifie (1) – (4) est une fonction de répartition d'une variable aléatoire sur Ω .

Pour une variable aléatoire discrète X , on peut définir une fonction de masse f_X . Celle ci est déterminée par la fonction de répartition F_X via :

$$f_X(x) = \lim_{n \rightarrow +\infty} \frac{F_X(x) - F_X(x - \frac{1}{n})}{(n+1) - n}.$$

Elle apparaît comme une version discrète de dérivée à droite en chaque point. Si on reverse le point de vue, la fonction de répartition peut être interprétée comme une primitive de la fonction de masse. Ceci suggère la définition suivante.

5. $\mathcal{B}(\mathbb{R})$ désigne la tribu des boréliens.

DÉFINITION II.3.3. Une variable aléatoire $X : \Omega \rightarrow \mathbb{R}$ est une variable aléatoire à densité s'il existe une fonction $f_X \in \mathcal{C}_m(\mathbb{R}, \mathbb{R})$ positive telle que⁶ pour tout $x \in \mathbb{R}$,

$$F_X(x) = \int_{-\infty}^x f(t) dt.$$

La fonction f_X est alors appelée une densité de probabilité de X .

Notez que pour être une densité, f_X doit vérifier la normalisation suivante :

$$\int_{\mathbb{R}} f(t) dt = 1.$$

REMARQUE II.3.4. La densité de probabilité n'est pas unique : deux telles fonctions diffèrent sur "un ensemble de mesure nulle". Par exemple, si on modifie une densité de probabilité en ses points de discontinuité, on ne change pas la fonction de répartition.

3.2. Propriétés, moments et indépendance. Les fonctions de répartition des variables aléatoires à densité ont le bon goût d'être continues, et même plus :

PROPOSITION II.3.5. Soit F_X la fonction de répartition d'une variable aléatoire à densité X de densité f_X . Alors $F_X \in \mathcal{C}_m^1(\mathbb{R}, \mathbb{R})$. Plus précisément, F_X est continue sur \mathbb{R} et dérivable en dehors de l'ensemble E des points de discontinuité de f_X , avec pour $x \in \mathbb{R} \setminus E$:

$$F_X'(x) = f_X(x).$$

REMARQUE II.3.6. Si X est une variable aléatoire à densité, on dit que \mathbb{P}_X est *absolument* continue (par rapport à la mesure de Lebesgue). Notez qu'il existe des fonctions de répartition (assez pathologiques on en convient) continues mais qui ne sont pas associées à des mesures de probabilités absolument continues. Les mesures correspondantes sont dites *singulières*.

Dans la suite, X désigne une variable aléatoire à densité, de densité f_X . On peut également retrouver la loi de X à partir de sa densité f_X :

LEMME II.3.7. Pour tout $[a, b] \subset \mathbb{R}$, on a

$$\mathbb{P}_X([a, b]) = \int_a^b f_X(t) dt.$$

On donne des exemples de variables aléatoires à densité, ainsi que certaines de leurs propriétés, dans l'Annexe D.

On peut définir l'indépendance de variables aléatoires en toute généralité. On utilisera ici la définition suivante⁷ :

DÉFINITION II.3.8. Deux variables aléatoires X et Y sur $(\Omega, \mathcal{J}, \mathbb{P})$ sont *indépendantes* si pour tout $a, b \in \mathbb{R}^2$, les événements $\{X \leq a\}$ et $\{Y \leq b\}$ sont indépendants. Une famille au plus dénombrable $(X_i)_{i \in I}$ de variables aléatoires discrètes est indépendante si les événements $\{X_j \leq a_j\}_{j \in J}$ sont indépendants pour tout $a_j \in \mathbb{R}$, $j \in J$ et $J \subset I$ fini.

Comme dans le cas discret, on définit les moments :

6. Le lecteur doit être averti que cette définition n'est pas la plus générale. Avec la théorie de la mesure en main, on suppose simplement f mesurable et positive.

7. Cette définition est bien équivalente à celle donnée pour les variables aléatoires discrètes. La définition générale standard fait intervenir les boréliens, que nous ne discuteront pas ici.

DÉFINITION II.3.9. Supposons que

$$\int_{\mathbb{R}} |x|f_X(x) dx < +\infty.$$

On définit l'espérance de X comme étant

$$\mathbb{E}(X) = \int_{\mathbb{R}} xf_X(x) dx.$$

L'espérance est linéaire. Une preuve naturelle repose sur la théorie de l'intégration de Lebesgue sur $(\Omega, \mathcal{T}, \mathbb{P})$ et ne sera pas abordée. Cette théorie permet de définir l'espérance de toute variable aléatoire "intégrable". On peut également, à l'aide de cette théorie, démontrer le "théorème de transfert" :

PROPOSITION II.3.10 (Théorème de transfert). Soit $\phi \in \mathcal{C}_m(\mathbb{R}, \mathbb{R})$. On suppose que

$$\int_{\mathbb{R}} |\phi(s)|f_X(s) ds < +\infty.$$

Alors l'espérance de la variable aléatoire $\phi(X)$ vérifie

$$\mathbb{E}(\phi(X)) = \int_{\mathbb{R}} \phi(s)f_X(s)ds.$$

Cette proposition permet alors de définir les moments d'ordre supérieur d'une variable aléatoire à densité :

DÉFINITION II.3.11. Soit $n \in \mathbb{N}^*$. Si

$$\int_{\mathbb{R}} |x^n|f_X(x) dx < +\infty$$

on appelle $\mathbb{E}(X^n)$ le moment d'ordre n de X . La variance de X est définie par

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2.$$

REMARQUE II.3.12. Les propriétés de l'espérance et de la variance listées dans la Section 2.2, notamment les inégalités de Bienaymé-Tchebychev et de Jensen, sont encore valables pour les variables aléatoires à densité (sous réserve de convergence des intégrales considérées).

4. Convergences de variables aléatoires

On termine ce cours par deux résultats fondamentaux sur la convergence de variables aléatoires. Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. On s'intéresse à une suite de réalisations d'une même expérience aléatoire, et on aimerait comprendre le comportement asymptotique d'une variable aléatoire que l'on observe à chaque réalisation. Idéalement, on suppose que les observations sont indépendantes. Ceci est modélisé par une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ indépendantes et identiquement distribuées (i.i.d.), c'est à dire de même loi. On admettra l'existence de telles suites.

THÉORÈME II.4.1 (Loi faible des grands nombres). Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes et de même loi, admettant une espérance μ et une variance finie. Soit, pour $n \in \mathbb{N}^*$,

$$S_n := \frac{1}{n} \sum_{k=1}^n X_k.$$

Alors pour tout $\epsilon > 0$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|S_n - \mu| \geq \epsilon) = 0$$

Ce théorème stipule que la moyenne de ces variables aléatoires converge *en probabilité*⁸ vers leur espérance.

Le deuxième résultat que l'on va énoncer précise la loi vers laquelle la moyenne de variables aléatoires i.i.d. converge. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires i.i.d. qui admettent une espérance μ et une variance $0 < \sigma^2 < +\infty$. On pose pour $n \in \mathbb{N}^*$,

$$S_n := \frac{1}{n} \sum_{k=1}^n X_k.$$

La moyenne S_n est alors d'espérance μ et d'écart type $\frac{\sigma}{\sqrt{n}}$. On renormalise cette moyenne en introduisant la variable aléatoire *réduite et centrée* :

$$Z_n := \sqrt{n} \frac{(S_n - \mu)}{\sigma}$$

qui est d'espérance nulle et d'écart-type 1. Le théorème central limite nous dit alors que $(Z_n)_{n \in \mathbb{N}^*}$ converge *en loi* vers une loi normale standard (i.e. une loi normale de moyenne nulle et d'écart-type 1). Soit X une telle variable aléatoire :

$$\forall x \in \mathbb{R}, \mathbb{P}(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt.$$

THÉORÈME II.4.2 (Théorème central limite). *Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires i.i.d. qui admettent une espérance μ et une variance $0 < \sigma^2 < +\infty$. On pose pour tout $n \in \mathbb{N}^*$*

$$Z_n := \sqrt{n} \frac{\left(\frac{1}{n} \sum_{k=1}^n X_k\right) - \mu}{\sigma}.$$

Alors pour tout $x \in \mathbb{R}$,

$$\lim_{n \rightarrow +\infty} \mathbb{P}(Z_n \leq x) = \mathbb{P}(X \leq x).$$

Autrement dit, F_{Z_n} converge simplement vers F_X . Il est remarquable que cette limite soit valable *quelque soit* la loi de la variable aléatoire de départ.

8. La loi forte des grands nombres stipule que cette convergence est en fait *presque sûre*

Annexe A

Primitives usuelles

| Fonction | Primitive |
|--------------------------|-----------------------------|
| e^x | e^x |
| $\ln(x)$ | $x \ln(x) - x$ |
| $\frac{1}{x}$ | $\ln(x)$ |
| $x^a (a \neq -1)$ | $\frac{x^{a+1}}{a+1}$ |
| $\sin(x)$ | $-\cos(x)$ |
| $\cos(x)$ | $\sin(x)$ |
| $\tan(x)$ | $-\ln(\cos(x))$ |
| $\operatorname{sh}(x)$ | $\operatorname{ch}(x)$ |
| $\operatorname{ch}(x)$ | $\operatorname{sh}(x)$ |
| $\operatorname{th}(x)$ | $\ln(\operatorname{ch}(x))$ |
| $\frac{1}{1+x^2}$ | $\arctan(x)$ |
| $\frac{1}{\sqrt{1-x^2}}$ | $\arcsin(x)$ |

Annexe B

Intégrales généralisées types

Soient $a \in \mathbb{R}$, $b \in \mathbb{R}_+^*$, $\lambda \in \mathbb{R}$ et $\alpha \in \mathbb{R}$. Le tableau suivant liste des résultats de convergence classique pour des intégrales impropres :

| Intégrale | Convergente | Divergente |
|--|---------------|------------------|
| $\int_a^{+\infty} e^{\lambda x} dx$ | $\lambda < 0$ | $\lambda \geq 0$ |
| $\int_{-\infty}^a e^{\lambda x} dx$ | $\lambda > 0$ | $\lambda \leq 0$ |
| $\int_b^{+\infty} \frac{dx}{x^\alpha}$ | $\alpha > 1$ | $\alpha \leq 1$ |
| $\int_0^b \frac{dx}{x^\alpha}$ | $\alpha < 1$ | $\alpha \geq 1$ |
| $\int_b^{+\infty} \ln(x) dx$ | jamais | toujours |
| $\int_0^b \ln(x) dx$ | toujours | jamais |

Annexe C

Variables aléatoires discrètes usuelles

Le tableau suivant liste quelques variables aléatoires discrètes usuelles, ainsi que leurs espérances et variances :

| Variable aléatoire (paramètres) | fonction de masse | Espérance | Variance |
|---|---|----------------------|------------------------|
| Uniforme ($\llbracket a, b \rrbracket$) | $f_X(k) = \frac{1}{n}, n = b - a + 1$ | $\frac{b + a}{2}$ | $\frac{n^2 - 1}{12}$ |
| Bernoulli (p) | $f_X(x) = \begin{cases} p & \text{si } x = 1 \\ (1 - p) & \text{si } x = 0 \end{cases}$ | p | $p(1 - p)$ |
| Binomiale (n, p) | $f_X(k) = \binom{n}{k} p^k (1 - p)^{n-k}$ | np | $np(1 - p)$ |
| Poisson (λ) | $f_X(k) = \frac{\lambda^k}{k!} e^{-\lambda}, k \in \mathbb{N}$ | λ | λ |
| Géométrique (p) | $f_X(k) = p(1 - p)^{k-1}, k \in \mathbb{N}^*$ | $\frac{1}{p}$ | $\frac{1 - p}{p^2}$ |
| Pascal (r, p) | $f_X(k) = \binom{k + r - 1}{k} p^r (1 - p)^k$ | $\frac{r(1 - p)}{p}$ | $\frac{r(1 - p)}{p^2}$ |

Annexe D

Variables aléatoires à densité usuelles

Le tableau suivant liste quelques variables aléatoires à densité usuelles, ainsi que leurs espérances et variances :

| Variable aléatoire (paramètres) | densité | Espérance | Variance |
|---------------------------------|--|---------------------|-----------------------|
| Uniforme ($[a, b]$) | $f_X(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x)$ | $\frac{b+a}{2}$ | $\frac{(b-a)^2}{12}$ |
| Exponentielle ($\lambda > 0$) | $f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x)$ | λ^{-1} | λ^{-2} |
| Normale (μ, σ^2) | $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ | μ | σ^2 |
| Gamma ($\lambda > 0, t > 0$) | $f_X(x) = \frac{1}{\Gamma(t)} \lambda^t x^{t-1} e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x)$ | $\frac{t}{\lambda}$ | $\frac{t}{\lambda^2}$ |
| Cauchy | $f_X(x) = \frac{1}{\pi(1+x^2)}$ | $+\infty$ | non définie |

On rappelle que la fonction Γ est définie par

$$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx.$$

Bibliographie

- [B] Christophe BERTAULT, notes de cours en ligne <http://christophebertault.fr/cours-et-exercices/>
- [C] Loren COQUILLE, *Introduction aux probabilités*, <https://www-fourier.ujf-grenoble.fr/~coquilll/files/MAT243.pdf>
- [Ex7] Site exo7, cours et exercices en ligne
- [RW1] Xavier BUFF, Emmanuel HARLBERSTADT, François MOULIN, Monique RAMIS et Jacques SAULOY, *Mathématiques, tout-en-un pour la Licence 1*, sous la direction de J.-P. Ramis et A. Warusfel. Niveau L1. Dunod, 2007.
- [RW2] Xavier BUFF, Emmanuel HARLBERSTADT, François MOULIN, Monique RAMIS et Jacques SAULOY, *Mathématiques, tout-en-un pour la Licence 2*, sous la direction de J.-P. Ramis et A. Warusfel. Niveau L2. Dunod, 2007.