PROBABILITÉS

JEAN-PAUL CALVI

Références. Le livre de Woodroofe [2] est très riche et ne nécessite que des connaissances élémentaires. Celui de Williams [1] utilise le langage moderne (et dès que l'on quitte le niveau élémentaire, indispensable) de la théorie de la mesure.

---

1. Probabilités et modélisation des phénomènes aléatoires

1.1. **Espaces de probabilité.** Un triplet $(E,\mathcal{E},p)$ où E est un ensemble, $\mathcal{E} \subset \mathcal{P}(E)$ et p une une application de $\mathcal{E}$ dans $[0,1]$ est un **espace de probabilité** si

(1) $\mathcal{E}$ satisfait les propriétés suivantes :
   (a) Il contient l'ensemble vide et l'ensemble total : $\emptyset \in \mathcal{E}$, $E \in \mathcal{E}$,
   (b) Il est *stable par complémentation* : $A \in \mathcal{E} \Rightarrow E\backslash A \in \mathcal{E}$,
   (c) Il est *stable par réunion dénombrable* : $(A_i \in \mathcal{E}, i \in \mathbb{N}) \Rightarrow \cup_{i\in\mathbb{N}} A_i \in \mathcal{E}$.

(2) La fonction p vérifie
   (a) $p(E) = 1$,
   (b) Si $A_i \in \mathcal{E}$ pour $i \in \mathbb{N}$ et si $A_i \cap A_j = \emptyset$ pour $i \neq j$ alors
$$p(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} p(A_i).$$
   On dit que p est σ**-additive**.

L'ensemble $\mathcal{E}$ s'appelle une σ**-algèbre**, l'application p, une **probabilité** (ou une **mesure de probabilité**). En joignant les conditions (1b) et (1c), on obtient que $\mathcal{E}$ est *stable par intersection dénombrable.* Les éléments de $\mathcal{E}$ sont appelés **évènement**. L'ensemble E est parfois appelé **espace des échantillons** ou **univers**.

En prenant $A_0 = E$ et $A_i = \emptyset$ pour $i > 0$, on déduit de (2b) que $p(\emptyset) = 0$. Ensuite prenant $A_i = \emptyset$ à partir de $i = n+1$ on trouve la version "finie" : si $A_i \in \mathcal{E}$ pour $1 \leq i \leq n$ et si $A_i \cap A_j = \emptyset$ pour $i \neq j$ alors
$$p(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} p(A_i).$$

Les propriétés suivantes découlent (le vérifier) de la définition de manière similaire :

Proposition 1.1. *Soit $(E,\mathcal{E},p)$ un espace de probabilité.*

(1) *Si A et B sont dans $\mathcal{E}$ et si $A \subset B$ alors*
$$p(B\backslash A) = p(B) - p(A).$$
En particulier $p(A) \leq p(B)$ et $p(E\backslash A) = 1 - p(A)$.

Date: Novembre 2004.

Préparation au concours interne de l'agrégation de mathématiques. Merci de signaler les erreurs et les imprécisions en écrivant à calvi@picard.ups-tlse.fr.

(2) *Pour tous A et B dans $\mathcal{E}$ on a*
$$p(A \cup B) = p(A) + p(B) - p(A \cap B).$$
(3) *Si A_n est une suite croissante d'évènements (ie $A_n \subset A_{n+1}$) alors on a*
$$p(\cup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} p(A_n).$$
Similairement, si B_n est une suite décroissante d'évènements (ie $B_n \supset B_{n+1}$) alors on a
$$p(\cap_{n\in\mathbb{N}} B_n) = \lim_{n\to\infty} p(B_n).$$

Du dernier point on déduit aisément le résultat suivant. On dit qu'un évènement F est presque sûr lorsque $p(F) = 1$. Montrer que si (F_n) est une suite d'évènements presque sûrs alors $\cap_{n\in\mathbb{N}} F_n$ est aussi un évènement presque sûr.

Exercice 1.1. Soit $A_i \in \mathcal{E}$, $i = 1,\dots,n$. Démontrer que
$$p(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{J\in\binom{\mathbf{n}}{k}} p(\cap_{j\in J} A_j)$$
où $\binom{\mathbf{n}}{k}$ désigne l'ensemble des parties à k éléments dans $\mathbf{n} = \{1,2,\dots,n\}$. On pourra utiliser une récurrence sur n. Cette formule importante est connue sous le nom de **formule du crible**. L'exercice 1.4 ci-dessous est une application de cette formule.

1.2. **Exemples fondamentaux.**

1.2.1. *Le cas fini.*

Théorème 1.1. *Soit E un ensemble fini et f une fonction positive définie sur E telle que $\sum_{e\in E} f(e) = 1$. Alors $(E,\mathcal{P}(E),p)$ est un espace de probabilité lorsque p est défini pour tout $A \in \mathcal{P}(E)$ par $p(A) = \sum_{e\in A} f(e)$.*

Comment. Lorsque $\text{card}(E) = n$ et $f(e) = 1/n$ pour tout $e \in E$ on retrouve $p(A) = \text{card}(A)/n$ qui est l'élémentaire "nombre de cas favorables / nombre total de cas". La théorie des probabilités sur un ensemble fini se réduit essentiellement au dénombrement d'ensembles finis. Elle est une sous-partie de la combinatoire.

1.2.2. *Modélisation d'un phénomène élémentaire.* Une expérience qui a deux issues possibles, disons P et F, est modélisée par l'espace de probabilité $(E,\mathcal{E},p)$ ou $E = \{P,F\}$, $\mathcal{E} = \mathcal{P}(E) = \{\emptyset,\{F\},\{P\},E\}$ et $p(P) = \alpha_1$, $p(F) = \alpha_2$ avec $\alpha_1 + \alpha_2 = 1$.

La fixation des constantes α_1 et α_2 n'est pas un travail mathématique, elle provient de l'expérience, éventuellement de l'intuition.

Si la même expérience est répétée n fois, le modèle, déduit du précédent, sera $(E^n,\mathcal{P}(E^n),\overline{p})$ où
$$\overline{p}(\{(\Box_1,\Box_2,\dots,\Box_n)\}) = p(\{\Box_1\}).p(\{\Box_2\})\dots p(\{\Box_n\})\}$$
avec $\Box_i \in \{P,F\}$.

Exercice 1.2. Dans l'espace $(E^n,\mathcal{P}(E^n),\overline{p})$ ci-dessus. Soit $1 \le r \le k \le n$. Calculer la probabilité de l'évènement suivant : le r-ième P apparaît à la k-ième expérience.

Exercice 1.3. Donner une modélisation d'une expérience aléatoire à 3 issues répétées n fois.

Exercice 1.4. L'ordinateur d'une banque envoie chaque mois un relevé de compte aux n clients de la banque. Une erreur de programmation fait mélanger aléatoirement les relevés de compte et les enveloppes (qui portent l'adresse du destinataire). Construire un espace de probabilité qui modélisent ce phénomène puis calculer la probabilité qu'au moins un client reçoive son propre relevé de compte et montrer que lorsque n est grand cette probabilité est proche de 0,63.

1.2.3. *Le cas infini dénombrable.*

Théorème 1.2. *Soient $E = \mathbb{N}$ (ou plus généralement un quelconque ensemble infini dénombrable) et (p_n) une suite de nombres positifs telle que $\sum_{n=0}^{\infty} p_n = 1$. Le triplet $(\mathbb{N},\mathcal{P}(\mathbb{N}),p)$ est un espace de probabilité lorsqu'on définit p par $p(A) = \sum_{n\in A} p_n$, $A \subset \mathbb{N}$.*

Notons que les nombres p_n étant positifs, dans le cas où A est dénombrable, $A = \{n_i : i \in \mathbb{N}\}$ la quantité $p(A) = \sum_{n\in A} p_n$ est égal à la somme de la série $\sum_{i=0}^{\infty} p_{n_i}$.

Exercice 1.5. Soit $\lambda > 0$. Pour tout $n \in \mathbb{N}$ on définit $p(\{n\}) = \frac{1}{n!}\lambda^n e^{-\lambda}$. Montrer que le triplet $(\mathbb{N},\mathcal{P}(\mathbb{N}),p)$ est un espace de probabilité.

On dit que la probabilité p suit une **loi de Poisson** de paramètre λ, elle-même appelée **loi des évènements rares** car $p(n)$ devient très petit lorsque n est grand. En pratique, elle est surtout une loi d'approximation (de la loi de Bernouilli, voir plus loin). Typiquement, $p(n)$ est la probabilité (l'approximation de la probabilité) que, parmi les N objets fabriqués par une machine, n soient défectueux avec un rapport n/N très proche de 0.

1.2.4. *Le cas continu.* Soit $E = I$ un intervalle de $\mathbb{R}$, pour $\mathcal{E}$ on prend la σ-algèbre des boréliens de I, c'est-à-dire la plus petite σ-algèbre contenant tous les sous-intervalles de I. Elle est notée $\mathcal{B}(I)$. Cette σ-algèbre est très vaste. Il faut même recourir à l'axiome du choix pour montrer l'existence d'ensembles qui ne sont pas boréliens. Etant donnée une fonction f continue positive telle que $\int_I f(x)dx = 1$, on démontre qu'il existe une unique probabilité p sur $(I,\mathcal{B}(I))$ telle que pour tout intervalle $J \in I$ on ait $p(J) = \int_J f(x)dx$. L'application f s'appelle la **densité** de la mesure de probabilité p. On notera $p = fdx$. On définit des espaces de probabilités similaires en faisant intervenir des ensembles et des fonctions de plusieurs variables. Ici les intégrales considérées sont des intégrales (éventuellement généralisées) de Riemann alors que le cadre naturel (qui permet l'intégration sur les boréliens quelconques) est celui de l'intégrale de Lebesgue que nous ne pouvons pas ici évoquer. En pratique, on se limitera à considérer des densités continues et des probabilités d'intervalles (ou de réunion finie d'intervalles) de sorte que le cadre inadapté de l'intégrale de Riemann n'induira pas d'inconvénient grave.

Exercice 1.6. Montrer que tout sous-ensemble fini de I appartient à $\mathcal{B}(I)$, puis que tout sous-ensemble de I infini dénombrable appartient à $\mathcal{B}(I)$.

1.2.5. *Modélisation d'un phénomène élémentaire.* Supposons qu'on choisisse un point aléatoirement dans l'intervalle $[0,1]$ (de telle sorte qu'aucun point ne soit privilégié par rapport aux autres). Le modèle sera $([0,1],\mathcal{B}([0,1]),dx)$ de sorte que $p([a,b]) = b - a$ pour tout $[a,b] \subset [0,1]$.

Exercice 1.7. Soit C le cercle unité de centre l'origine dans le plan munie de sa topologie habituelle (induite de celle du plan). On définit $\mathcal{B}(C)$ comme la plus petite σ-algèbre contenant les ouverts de C. Donner un exemple de phénomène aléatoire dont la modélisation serait l'espace de probabilité $(C,\mathcal{B}(C),p)$ où p est telle que pour tout arc $J \subset C$ on ait $p(J) = (2\pi)^{-1}\text{longueur}(J)$.

1.2.6. *Loi normale.* L'espace de probabilité $(\mathbb{R},\mathcal{B}(\mathbb{R}),p)$ où

$$p([a,b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

est l'un des plus importants de toute la théorie des probabilités. On note $p = \mathbf{N}(0,1)$. La lettre N est mise pour "Normale". Plus généralement, pour $(\mu,\sigma) \in (\mathbb{R},\mathbb{R}^{\star+})$, on considère $p = \mathbf{N}(\mu,\sigma)$ définie par

$$p([a,b]) = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

2. Constructions de nouveaux espaces de probabilités à partir d'espaces donnés

2.1. **Probabilités conditionnelles.**

2.1.1. *Définition.* Soient $(E,\mathcal{E},p)$ un espace de probabilité et $B \in \mathcal{E}$. L'ensemble

$$\mathcal{E}_B := \{A \cap B, A \in \mathcal{E}\} = \{C \in \mathcal{E} : C \subset B\}$$

est une σ-algèbre. La **probabilité de** $A \in \mathcal{E}$ **sachant** B est définie par

$$p_B(A) = \frac{p(A \cap B)}{p(B)}.$$

Théorème 2.1. *Le triplet $(E,\mathcal{E},p_B)$ est un espace de probabilité.*

La probabilité p_B s'appelle une **probabilité conditionnelle**. On emploie souvent la notation $p_B(A) = p(A|B)$.

2.1.2. *Formules de Bayes.*

Proposition 2.1. *Soit $(E,\mathcal{E},p)$ un espace de probabilité.*

(1) *Soient $A_1, \dots A_n$ n évènements de $\mathcal{E}$. On pose $B_k = \cap_{1\le j\le k} A_j$, $k = 1, \dots, n$. Si $p(B_{n-1}) > 0$ alors on a*

$$p(B_n) = p(B_1) \prod_{k=2}^{n} p(B_k|B_{k-1}).$$

(2) *Soient (B_n) une suite (finie ou non) d'évènements de $\mathcal{E}$ de probabilité non nulle. On suppose que ces évènements sont deux à deux disjoints et qu'ils sont* ***exhaustifs*** *c'est-à-dire que leur réunion est égale à E. Pour tout $A \in \mathcal{E}$ on a*
$$p(A) = \sum_{n \geq 0} p(A|B_n)P(B_n).$$

(3) *Avec les mêmes hypothèses et si de plus $p(A) > 0$ on a*
$$p(B_j|A) = \frac{p(A|B_j)p(B_j)}{\sum_{n\geq 0} p(A|B_n)p(B_n)}.$$

Exercice 2.1. Etude d'un problème de transmission. Un signal (0 ou 1) est émis d'un poste A et passe par 4 relais avant d'être réceptionné par le poste B. La transmission est représentée par le schéma suivant
$$A \dashrightarrow \boxed{1} \rightarrow \boxed{2} \rightarrow \boxed{3} \rightarrow \boxed{4} \dashrightarrow B$$
Lorsque le relais $\boxed{i}$ reçoit un 0, 9 fois sur 10 en moyenne il renvoie un 0, (dans le cas restant il le transforme en 1) et lorsqu'il reçoit un 1, dans 8 cas sur 10 en moyenne il renvoie un 1 (dans les deux cas restants il le transforme en 0). Modéliser le phénomène et calculer la probabilité pour qu'un 1 (resp. un 0) soit reçu quand un 1 (resp. un 0) est envoyé. L'émetteur A envoie le mot 00110, le récepteur peut reconnaître le mot s'il comporte un nombre d'erreur ≤ 1. Construire le modèle permettant de donner une réponse à la question : quelle est la probabilité que le mot soit reconnu par le récepteur B ?

Quid s'il s'agit de considérer un réseau de d relais ?

Exercice 2.2. Le tableau suivant montre (en pourcentage) le choix des hommes et des femmes d'une communauté C en faveur des partis politiques A, B et C.

	Hommes	Femmes
Parti A	20	25
Parti B	10	15
Parti C	15	15

Donner une modélisation permettant de résoudre le problème suivant. Quelle est la probabilité qu'une personne prise au hasard dans la communauté soit un homme sachant qu'il vote pour le parti A ?

2.2. **Variables aléatoires.**

2.2.1. *Définition.* Soit $(E,\mathcal{E},p)$ un espace de probabilité et $\mathbf{X}$ une fonction de E dans $\mathbb{R}$. On dit que $\mathbf{X}$ est une **variable aléatoire (v. a.)** sur E (ou plus précisément sur $(E,\mathcal{E})$) lorsque pour tout $W \in \mathcal{B}(\mathbb{R})$ on a $\mathbf{X}^{-1}(W) \in \mathcal{E}$. Autrement dit, les images réciproques des boréliens doivent être des évènements de $\mathcal{E}$. Les v.a. sont aussi appelées (dans un autre contexte) **fonctions $\mathcal{E}$-mesurables**. Lorsque $\mathcal{E} = \mathcal{P}(E)$ toute fonction $\mathbf{X} : E \to \mathbb{R}$ est une v. a.

Proposition 2.2. *Pour que $\mathbf{X}$ soit une v.a. il faut et il suffit que pour tout invervalle $I \subset \mathbb{R}$ on ait $\mathbf{X}^{-1}(I) \in \mathcal{E}$.*

Exercice 2.3. Démontrer la proposition. (On pourra montrer que l'ensemble $\mathcal{A}$ formé par les $W \in \mathcal{B}(\mathbb{R})$ tels que $X^{-1}(W) \in \mathcal{E}$ forment une σ-algèbre.)

Exercice 2.4. Soit $f : \mathbb{R} \to \mathbb{R}$. Montrer que si f est continue (sur $\mathbb{R}$) alors elle est une v.a. sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. Donner des exemples de v.a. qui ne soient pas continues.

2.2.2. *Notations.* Traditionnellement, l'évènement $\mathbf{X}^{-1}(W)$ est (presque toujours) noté $\mathbf{X} \in W$ de sorte que l'on pourra calculer $p(\mathbf{X} \in W)$. L'évènement $\mathbf{X} \in \{a\}$ est encore simplifié en $\mathbf{X} = a$.

Comment. Ces notations — et le fait d'appeler "variable" une fonction — sont certainement assez malheureux et expliquent en partie (avec d'autres bizareries de langage) les difficultés qu'ont beaucoup d'étudiants à assimiler le langage probabiliste. Ceci étant dit, elles ne seront jamais changées et il vaut mieux en prendre rapidement son parti.

2.2.3. *Variable aléatoire et probabilités.* Soient $(E,\mathcal{E},p)$ un espace de probabilité et $\mathbf{X}$ une v.a. sur $(E,\mathcal{E})$. On définit la fonction $p_{\mathbf{X}}$ sur $\mathcal{B}(\mathbb{R})$ par $p_{\mathbf{X}}(W) = p(\mathbf{X} \in W)$.

Théorème 2.2. *Le triplet* $(\mathbb{R},\mathcal{B}(\mathbb{R}),p_{\mathbf{X}})$ *est un espace de probabilité.*

La probabilité $p_{\mathbf{X}}$ s'appelle la **loi** de $\mathbf{X}$.

2.2.4. *Loi de Bernoulli.* Considérons l'espace de probabilité $(E^n,\mathcal{P}(E^n),\overline{p})$ où $E = \{P,F\}$. Pour $i = 0,\dots,n$ et $e = (e_1,\dots,e_n) \in E^n$ on définit $\mathbf{X}_i(e) = 1$ si $e_i = P$ et 0 sinon. La variable aléatoire

$$\mathbf{S}_n = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n$$

est particulièrement importante. Elle prend ses valeurs dans $\{0,1,\dots,n\}$ et "$\mathbf{S}_n = k$" est l'évènement correspondant à "parmi les n expériences, le résultat a été k fois P".

Théorème 2.3. *Pour* $k \in \{0,\dots,n\}$ *on a* $\overline{p}(\mathbf{S}_n = k) = \binom{n}{k}\alpha_1^k\alpha_2^{n-k}$.

La loi $p_{\mathbf{S}_n}$ s'appelle une **loi de Bernoulli** (et on dit que $\mathbf{X} = \mathbf{S}_n$ suit une loi de Bernoulli ou est une variable aléatoire de Bernoulli). Puisque la probabilité $\overline{p}$ dépend des paramètres n et $\alpha_1 = r$ (qui oblige $\alpha_2 = 1 - r$). On écrira, lorsqu'il sera nécessaire d'être précis, $\overline{p}_{n,r}$. On note aussi $\overline{p}_{n,r}(\mathbf{S}_n = k) = \mathbf{b}(k,n,r)$. On dit que le couple (n,r) est le paramètre de loi de Bernoulli.

Exercice 2.5. Soit (r_n) une suite de nombres dans $]0,1[$ telle que $\lim_{n\to\infty} nr_n = \lambda \in]0,\infty[$. Montrer que $\forall k \in \mathbb{N}$, $\lim_{n\to\infty} \overline{p}_{n,r_n}(S_n = k) = \frac{1}{k!}\lambda^k e^{-\lambda}$.

2.2.5. *Variables aléatoires discrètes.* On dit qu'une variable aléatoire $\mathbf{X}$ est **discrète** finie (*resp.* discrète infinie) s'il existe un ensemble $C \subset \mathbb{R}$ fini (*resp.* infini dénombrable) tel que $P(\mathbf{X} \in C) = 1$. Les variables aléatoires de Bernoulli sont discrètes finies.

2.2.6. *Variables aléatoires continues.* On dit qu'une variable aléatoire $\mathbf{X}$ est continue s'il existe une fonction continue $f = f_{\mathbf{X}}$ telle que $p_{\mathbf{X}} = f_{\mathbf{X}}dx$. L'application $f_{\mathbf{X}}$ s'appelle la densité de $\mathbf{X}$. Lorsque $p_{\mathbf{X}} = \mathbf{N}(\mu,\sigma)$ on dit que $\mathbf{X}$ suit une loi normale de paramètre (μ,σ).

2.2.7. *Structure de l'ensemble des v.a.*

Théorème 2.4. *Soient* $\mathbf{X}_1$ *et* $\mathbf{X}_2$ *deux variables aléatoires sur* $(E,\mathcal{E})$ *et* $\lambda \in \mathbb{R}$ *alors* $\mathbf{X}_1 + \mathbf{X}_2$, $\lambda \mathbf{X}_1$ *et* $\mathbf{X}_1\mathbf{X}_2$ *sont des v.a.*

Théorème 2.5. *Si* $f : \mathbb{R} \to \mathbb{R}$ *est un v.a. sur* $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *et si* $\mathbf{X}$ *est une v.a. sur* $(E,\mathcal{E})$ *alors* $f \circ \mathbf{X}$ *est un v.a sur* $(E,\mathcal{E})$.

On trouve très souvent la notation $f(\mathbf{X})$ pour $f \circ \mathbf{X}$.

2.3. **Probabilité produit.**

Théorème 2.6. *Soient* $(E_1,\mathcal{E}_1,p_1)$ *et* $(E_2,\mathcal{E}_2,p_2)$ *deux espaces de probabilités (non nécessairement distincts). Il existe une unique fonction* p *telle que*

(1) *Le triplet* $(E_1 \times E_2,\sigma(\mathcal{E}_1 \times \mathcal{E}_2),p)$ *— où* $\sigma(\mathcal{E}_1 \times \mathcal{E}_2)$ *est la plus petite* σ*-algèbre contenant* $\mathcal{E}_1 \times \mathcal{E}_2 = \{A \times B, A \in \mathcal{E}_1, B \in \mathcal{E}_2\}$ *— est un espace de probabilité.*
(2) *Pour tous* $A \in \mathcal{E}_1$ *et* $B \in \mathcal{E}_1$ *on a* $p(A \times B) = p_1(A).p_2(B)$.

L'espace $(E_1 \times E_2,\sigma(\mathcal{E}_1 \times \mathcal{E}_2),p)$ s'appelle **l'espace produit** de $(E_1,\mathcal{E}_1,p_1)$ par $(E_2,\mathcal{E}_2,p_2)$. En particulier, p s'appelle la **probabilité produit** de p_1 par p_2 et est notée $p = p_1 \otimes p_2$. L'existence et l'unicité d'une telle probabilité est aisément vérifiée dans le cas fini et infini dénombrable (où $\sigma(\mathcal{E}_1 \times \mathcal{E}_2) = \mathcal{P}(E_1 \times E_2)$). Elle est beaucoup plus compliquée dans le cas continu. Ces espaces produits servent à modéliser un ensemble de phénomènes, souvent un même phénomène répété plusieurs fois comme on l'a vu dans les exemples précédents.

Si $(E_1,\mathcal{E}_1,p_1) = (I,\mathcal{B}(I),f_1dx)$ et $(E_2,\mathcal{E}_2,p_2) = (J,\mathcal{B}(J),f_2dx)$ où I (resp. J) est un intervalle de $\mathbb{R}$ et f_1 (resp. f_2) une fonction continue sur I (resp. J) alors pour tout ensemble C quarrable dans $I \times J$ (en réalité pour tout borélien de $I \times J$) on a

$$(p_1 \otimes p_2)(C) = \iint_C f_1(x)f_2(y)dxdy.$$

3. Indépendance

3.1. **Définition.** Soit $(E,\mathcal{E},p)$ un espace de probabilité. On dit que deux évènements A et B dans $\mathcal{E}$ sont **indépendants** lorsque

$$p(A \cap B) = p(A).p(B).$$

Plus généralement, n évènements $A_1,A_2,\ldots,A_n$ sont dit indépendants (ou plus précisément **deux à deux indépendants**) lorsque pour tous $i \neq j$ on a

$$p(A_i \cap A_j) = p(A_i).p(A_j).$$

Il est très facile de trouver des exemples d'évènements indépendants dans les espaces de probabilité produit. Si $p = p_1 \otimes p_2$ sur $(E_1 \times E_2,\sigma(\mathcal{E}_1 \times \mathcal{E}_2))$ et si $A = E_1 \times L$ et $B = T \times E_2$ avec $L \in \mathcal{E}_2$ et $T \in \mathcal{E}_1$ alors les évènements A et L sont indépendants. En effet, $A \cap B = T \times L$ et par définition de la probabilité produit $p(T \times L) = p_1(T).p_2(L) = p(T \times E_2).p(E_1 \times L) = p(A).p(B)$.

Exercice 3.1. Soient $A_1,A_2,\ldots A_n$ et $B_1,B_2,\ldots B_n$ deux familles de n évènements deux à deux disjoints dans $\mathcal{E}$. On suppose que pour tout (i,j) les évènements A_i et B_j sont indépendants. Montrer que $A = \cup_{i=1}^n A_i$ et $B = \cup_{j=1}^n B_j$ sont indépendants.

3.2. **Indépendance des variables aléatoires.** Deux variables aléatoires $\mathbf{X}_1$ et $\mathbf{X}_2$ sur $(E,\mathcal{E})$ sont dites indépendantes si, pour tous W et W' dans $\mathcal{B}(\mathbb{R})$, les évènements $\mathbf{X}_1 \in W$ et $\mathbf{X}_2 \in W'$ sont indépendants. La définition s'étend immédiatement au cas de n variables aléatoires.

Exercice 3.2. Montrer que pour que $\mathbf{X}_1$ et $\mathbf{X}_2$ sur $(E,\mathcal{E})$ soient indépendantes il faut et il suffit que pour tous les intervalles I et J dans $\mathbb{R}$, on ait a

$$p(\{\mathbf{X}_1 \in I\} \cap \{\mathbf{X}_1 \in J\}) = p(\mathbf{X}_1 \in I).p(\mathbf{X}_2 \in J).$$

Théorème 3.1. *Soient $(E,\mathcal{E},p)$ un espace de probabilité et $\mathbf{X}_1$, $\mathbf{X}_2$ deux variables aléatoires indépendantes sur $(E,\mathcal{E})$ de densité (continue) respective f_1 et f_2. Pour tout C (quarrable) dans $\mathbb{R}^2$ on a*

$$(\mathbf{X}_1,\mathbf{X}_2) \in C = \{(x,y) \in E \times E : ((\mathbf{X}_1(x),\mathbf{X}_2(y)) \in C\} \in \mathcal{E}$$

et

$$p((\mathbf{X}_1,\mathbf{X}_2) \in C) = \iint_C f_1(x)f_2(y)dxdy.$$

3.3. **Somme de deux variables aléatoires.**

Théorème 3.2. *Soient $(E,\mathcal{E},p)$ un espace de probabilité et $\mathbf{X}_i$ $(i = 1,2)$ deux variables aléatoires indépendantes sur $(E,\mathcal{E})$. On étudie la variable aléatoire $\mathbf{X}_1 + \mathbf{X}_2$ sur $(E,\mathcal{E},p)$.*

(1) *Si $\mathbf{X}_i$ suit une loi de Bernoulli de paramètre (n_i,r), $i = \{1,2\}$, alors $\mathbf{X}_1+\mathbf{X}_2$ suit une loi de Bernoulli de paramètre $(n_1 + n_2,r)$.*
(2) *Si $\mathbf{X}_i$ suit une loi de Poisson de paramètre λ_i, $i = \{1,2\}$, alors $\mathbf{X}_1 + \mathbf{X}_2$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.*
(3) *Si $\mathbf{X}_i$ suit une loi normale de paramètre (μ_i,σ_i), $i = \{1,2\}$, alors $\mathbf{X}_1 + \mathbf{X}_2$ suit une loi normale de paramètre $(\mu_1 + \mu_2,\sqrt{\sigma_1^2 + \sigma_2^2})$.*

Exercice 3.3. Démontrer le théorème.

---

4. Espérance et notions associées

4.1. **Espérance.** L'absence de théorie de la mesure commence ici à poser problème. Nous serons obligés de donner 3 définitions différentes de ce qui est en réalité un seul et même objet mathématique. Soit $\mathbf{X}$ une variable aléatoire sur $(E,\mathcal{E},p)$.

4.1.1. *Le cas discret.* Si $\mathbf{X}$ est discrète finie alors **l'espérance** de $\mathbf{X}$, noté $E(\mathbf{X})$ est définie par

$$E(\mathbf{X}) = \sum_{i\in C} ip(\mathbf{X} = i)$$

où C fini est tel que $p(\mathbf{X} \in C) = 1$.

Exercice 4.1. Déterminer l'espérance d'une loi de Bernoulli de paramètre (n,r).

4.1.2. *Le cas infini dénombrable.* Si $\mathbf{X}$ est discrète infinie alors l'espérance de $\mathbf{X}$, noté $E(\mathbf{X})$ est définie, *sous réserve que la série soit absolument convergente*, par

$$E(\mathbf{X}) = \sum_{n \in C} np(\mathbf{X} = n)$$

où C dénombrable est tel que $p(\mathbf{X} \in C) = 1$.

Exercice 4.2. Déterminer l'espérance d'une variable aléatoire suivant une loi de Poisson (ie telle que $p(\mathbf{X} = n) = \lambda^n e^{-\lambda}/n!$ pour $n \in \mathbb{N}$.

4.1.3. *Le cas continu.* Si $\mathbf{X}$ est une variable aléatoire continue de densité f alors l'espérance de $\mathbf{X}$, noté $E(\mathbf{X})$ est définie, *sous réserve que l'intégrale soit absolument convergente*, par

$$E(\mathbf{X}) = \int_{-\infty}^{\infty} xf(x)dx$$

Exercice 4.3. Déterminer l'espérance d'une variable aléatoire suivant une loi normale de paramètre (μ,σ).

4.1.4. *Extension de la définition.* Si $\mathbf{X}$ est une somme de variables aléatoires $\mathbf{X} = \mathbf{X_1} + \mathbf{X_2} + \cdots + \mathbf{X_n}$ ou chacun des termes $\mathbf{X}_i$ est d'une des trois formes ci-dessus, on définit l'espérance de $\mathbf{X}$ comme la somme des espérances de $\mathbf{X}_i$ sous réserve que chacune de ces n espérances soit bien définie.

4.2. Propriétés de l'espérance.

4.2.1. *Linéarité.*

Théorème 4.1. *Soient $\mathbf{X}_1$ et $\mathbf{X}_2$ deux variables aléatoires sur un même espace de probabilité $(E,\mathcal{E},p)$ et $(\alpha,\beta) \in \mathbb{R}$. On a*

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

De plus, si $p(\mathbf{X} \geq \mathbf{Y}) = 1$ alors $E(\mathbf{X}) \geq E(\mathbf{Y})$.

4.2.2. *Espérance d'une image.*

Théorème 4.2. *Soient $\mathbf{X}$ une variable aléatoire sur $(E,\mathcal{E},p)$ et h une variable aléatoire sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.*

1. *Si $\mathbf{X}$ est fini discrète alors $E(h(\mathbf{X})) = \sum_{i \in C} h(i)p(\mathbf{X} = i)$*
2. *Si $\mathbf{X}$ est discrète infinie alors $E(h(\mathbf{X})) = \sum_{n \in C} h(n)p(\mathbf{X} = n)$*
3. *Si $\mathbf{X}$ est continue de densité f alors $E(h(\mathbf{X})) = \int_{-\infty}^{\infty} h(x)f(x)dx$*

Exercice 4.4. Démontrer le théorème dans le cas d'une variable aléatoire discrète (finie).

4.2.3. *Espérance et indépendance.*

Théorème 4.3. *Soient* $\mathbf{X}$ *et* $\mathbf{Y}$ *deux variables aléatoires (sur un même espace de probabilité) indépendantes. On a*

$$E(\mathbf{XY}) = E(\mathbf{X})E(\mathbf{Y}).$$

Exercice 4.5. Démontrer le théorème dans le cas de deux variables aléatoires discrètes (finies).

4.3. **Variance.** Soi $\mathbf{X}$ une variable aléatoire sur $(E,\mathcal{E},p)$ d'espérance μ. La variance $\sigma^2 = \sigma^2(\mathbf{X})$ de $\mathbf{X}$ est définie, sous réserve de l'existence de l'espérance par

$$\sigma^2 = E((\mathbf{X} - \mu)^2) = E(\mathbf{X}^2) - \mu^2.$$

Le réel positif σ s'appelle l'écart type de $\mathbf{X}$.

Exercice 4.6. Montrer que $E((\mathbf{X} - a)^2)$ est minimisé lorsque $a = \mu$.

Théorème 4.4. *Soient* $\mathbf{X}$ *et* $\mathbf{Y}$ *deux variables aléatoires indépendantes. On a*

$$\sigma^2(\mathbf{X} + \mathbf{Y}) = \sigma^2(\mathbf{X}) + \sigma^2(\mathbf{Y}).$$

Exercice 4.7. Démontrer le théorème.

4.4. **Inégalités de Chebyshev.**

Théorème 4.5. *Soit* $\mathbf{X}$ *une variable aléatoire sur* $(E,\mathcal{E},p)$ *dont les espérances et variances sont bien définies. On pose* $\mu = E(\mathbf{X})$*. Pour tout* $c \geq 0$ *on a*

$$c^2 p(|\mathbf{X} - \mu| > c) \leq \sigma^2(\mathbf{X}).$$

Exercice 4.8. Démontrer le théorème.

Références

[1] D Williams. *Probability with Martingales.* Cambridge University Press, 1991.

[2] M Woodfroofe. *Probability with Applications.* Mc Graw-Hill, 1975.

Laboratoire de Mathématiques E. Picard, Université Paul Sabatier 31062 Toulouse Cedex FRANCE.

E-mail address: `calvi@picard.ups-tlse.fr`