COURS DE PROBABILITES 1

Licence 2 mathématiques

Semestre 4

N. CHENAVIER

Table des matières

Chapitre 1

Quelques notions pour les probabilités

Sommaire

1.1 Théorie des ensembles, dénombrement

Réunion, intersection, produit, complémentaire

Définition 1.1.1. *Soient A et B deux sous-ensembles d'un ensemble E.*

- *On appelle **intersection** de A et de B l'ensemble suivant :*

$$A \cap B = \{x : x \in A \text{ et } x \in B\}.$$

- *On appelle **réunion** de A et de B l'ensemble suivant :*

$$A \cup B = \{x : x \in A \text{ ou } x \in B\}.$$

Définition 1.1.2. *Soit $(A_i)_{i\in I}$ une famille de sous-ensembles d'un ensemble E.*

- *On appelle **intersection** des A_i, $i \in I$, l'ensemble suivant :*

$$\bigcap_{i\in I} A_i = \{x \in E : \forall i \in I, x \in A_i\}.$$

- *On appelle **réunion** des A_i, $i \in I$, l'ensemble suivant :*

$$\bigcup_{i\in I} A_i = \{x \in E : \exists i \in I, x \in A_i\}.$$

Définition 1.1.3. *Soient E un ensemble et A un sous-ensemble de E. On appelle **complémentaire** de A dans E l'ensemble suivant :*

$$A^c = \{x \in E : x \notin A\}.$$

Définition 1.1.4. *Soient A et B deux ensembles. On appelle **produit cartésien** de A et de B l'ensemble suivant :*

$$A \times B = \{(x, y) : x \in A \text{ et } y \in B\}.$$

Notation. *Soit E un ensemble. On désigne par $\mathcal{P}(E)$ l'ensemble des parties de E (c'est-à-dire de ses sous-ensembles). En d'autres termes,*

$$\mathcal{P}(E) = \{A : A \subset E\}.$$

Images directes et réciproques

Définition 1.1.5. *Soit f une application de E vers F et A un sous-ensemble de E. On appelle **image directe** de l'ensemble A par f l'ensemble, noté $f(A)$, des images des éléments de A par f. En d'autres termes,*

$$f(A) = \{f(x) : x \in A\}.$$

Définition 1.1.6. *Soit f une application de E vers F et B un sous-ensemble de F. On appelle **image réciproque** de l'ensemble B par f l'ensemble, noté $f^{-1}(B)$, des antécédants des éléments de B par f. En d'autres termes*

$$f^{-1}(B) = \{x \in E : f(x) \in B\}.$$

Dénombrement

Définition 1.1.7. *Soit E un ensemble ayant un nombre fini d'éléments. On appelle **cardinal** de E le nombre d'éléments de E. On note ce nombre $|E|$.*

Proposition 1.1.8. *(Cardinal d'une réunion, intersection) Soient A et B deux ensembles ayant un nombre fini d'éléments. Alors $|A \cup B| = |A| + |B| - |A \cap B|$.*

Proposition 1.1.9. *Soit (A_n) une suite de sous-ensembles deux à deux disjoints (cad $A_n \cap A_m = \emptyset$ dès lors que $n \neq m$) d'un ensemble E. Alors*

$$\left| \bigsqcup_{k=1}^{n} A_k \right| = \sum_{k=1}^{n} |A_k|.$$

Proposition 1.1.10. *(Cardinal d'un produit) Soient A et B deux ensembles ayant un nombre fini d'éléments. Alors $|A \times B| = |A| \times |B|$.*

Proposition 1.1.11. *(Cardinal du complémentaire) Si E est un ensemble ayant un nombre fini d'éléments, alors il en est de même pour tout sous-ensemble A. De plus, $|A^c| = |E| - |A|$.*

Proposition 1.1.12. *(Cardinal de l'ensemble des parties) Soit E un ensemble ayant un nombre fini d'éléments. Alors $|\mathcal{P}(E)| = 2^{|E|}$.*

Proposition 1.1.13. *Soit n un entier, E un ensemble à n éléments et $k \leq n$.*

(i) Le nombre de k-uplets ordonnés d'éléments distincts de E, c'est-à-dire de la forme $(x_1, x_2, \ldots, x_k)$, avec $x_i \in E$ et $x_i \neq x_j$ pour tout $i \neq j$, est égal à

$$A_n^k = n(n-1)\cdots(n-k+1).$$

(ii) Le nombre de sous-ensembles de E de cardinal k, c'est-à-dire de la forme $\{x_1, x_2, \ldots, x_k\}$, avec $x_i \in E$ et $x_i \neq x_j$ pour tout $i \neq j$, est égal à

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{(n-k)!k!}.$$

Définition 1.1.14. *Soit n un entier naturel et $0 \leq k \leq n$. La quantité A_n^k s'appelle le nombre **d'arrangements** et le terme $\binom{n}{k}$ s'appelle le **coefficient binomial** de k parmi n.*

Exemple 1.1.15. Prenons un ensemble à $n = 3$ éléments, disons $E = \{a, b, c\}$.

1. Le nombre de couples constitués d'éléments distincts est $A_4^2 = 6$. Les couples sont (a, b), (b, a), (a, c), (c, a), (a, d) et (d, a).
2. Le nombre de pairs de E est égal à $\binom{3}{2} = 3$. Les pairs sont $\{a, b\}$, $\{a, c\}$ et $\{b, c\}$.

Le coefficient binomial intervient naturellement dans des problèmes de dénombrement comme en témoigne l'exemple ci-dessous.

Exemple 1.1.16. On prélève simultanément trois boules dans une urne contenant quatre boules rouges, trois boules vertes et une boule noire.

- Le nombre de possibilités de tirer trois boules rouges est égal à $\binom{4}{3} = 4$.
- Le nombre de tirages avec deux boules rouges et une boule verte est égale à $\binom{4}{2} \times \binom{3}{1} = 6 \times 3 = 18$.

1.2 Dénombrabilité

On distingue deux types d'ensembles : les ensembles finis et les ensembles infinis. Au sein des ensembles infinis, on distingue deux sous-classes : les ensembles (infinis) dénombrables et les ensembles indénombrables. La théorie des ensembles dans laquelle on se place est celle de ZFC à laquelle on rajoute l'axiome de l'infini (pour assurer l'existence de *l'ensemble* des entiers naturels $\mathbf{N}$).

Définition 1.2.1. *Un ensemble E est infini (au sens de Dedekind) s'il existe une injection $j : \mathbf{N} \to E$.*

Définition 1.2.2. *Soit E un ensemble. On dit que E est **dénombrable** s'il s'injecte dans $\mathbf{N}$, c'est-à-dire s'il existe une application injective $i : E \to \mathbf{N}$.*

Tous les ensembles finis sont dénombrables. Le résultat ci-dessous affirme que si un ensemble est infini et dénombrable alors il est en bijection avec $\mathbf{N}$.

Proposition 1.2.3. *Soit E un ensemble infini. Les propriétés suivantes sont équivalentes :*

(i) E est dénombrable ;

(ii) il existe une bijection $\varphi : E \to \mathbf{N}$;

(iii) il existe une surjection $s : \mathbf{N} \to E$.

Informellement, un ensemble est dénombrable s'il est "suffisamment petit" pour que l'on puisse "numéroter" ces éléments. Plus précisément, un ensemble est dénombrable s'il est fini ou s'il peut s'écrire sous la forme $E = \{x_n\}_{n \in \mathbf{N}}$, où les x_i sont des éléments deux à deux différents.

Notation. *Soient E et F deux ensembles.*

- *Lorsque E s'injecte dans F, on note $E \hookrightarrow F$.*

- *Lorsqu'il existe une surjection de F dans E, on note $F \twoheadrightarrow E$.*
- *Lorsque E et F sont en bijection, on note $E \longleftrightarrow F$.*

Le résultat ci-dessus peut alors s'énoncer de la façon suivante : si $\mathbf{N} \hookrightarrow E$, alors les propriétés suivantes sont équivalentes :

(i) $E \hookrightarrow \mathbf{N}$;

(ii) $E \longleftrightarrow \mathbf{N}$;

(iii) $\mathbf{N} \twoheadrightarrow E$.

Remarque 1.2.4. Si E est un ensemble dénombrable, et si F s'injecte dans E, alors F est dénombrable.

Exemple 1.2.5. Les ensembles $\mathbf{N}$, $\mathbf{N}^*$, $\mathbf{Z}$, $\mathbf{N}^2$, $\mathbf{Q}$ sont dénombrables.

Contre-exemple 1.2.6. Les ensembles $\mathbf{R}$, $\mathbf{R} \setminus \mathbf{Q}$, $\{0,1\}^{\mathbf{N}}$, $\mathcal{P}(\mathbf{N})$ ne sont pas dénombrables.

Proposition 1.2.7. *(i) Soient E et F deux ensembles dénombrables. Alors l'ensemble $E \cup F$ est dénombrable.*

(ii) Plus généralement, soient (E_n) une suite d'ensembles dénombrables. Alors $\bigcup_{n\geq 0} E_n$ est dénombrable.

Proposition 1.2.8. *Soient E et F deux ensembles dénombrables. Alors $E \times F$ est dénombrable.*

1.3 Séries numériques

Soit (u_n) une suite de nombres réels ou complexes. On définit les sommes partielles par

$$S_n = \sum_{k=0}^{n} u_k.$$

La suite des sommes partielles (S_n) s'appelle la **série** de terme général u_n. On la note $\sum_{n\geq 0} u_n$.

Définition 1.3.1. *On dit que la série $\sum_{n\geq 0} u_n$ est :*

- ***convergente** si la limite $\lim_{n\to\infty} S_n$ existe, et on note alors $\sum_{n=0}^{\infty} u_n := \lim_{N\to\infty} \sum_{n=0}^{N} u_n$ cette limite ;*
- ***divergente** si elle n'est pas convergente ;*
- ***absolument convergente** si la série $\sum_{n\geq 0} |u_n|$ est convergente.*

Remarque 1.3.2. Si (u_n) est une suite à termes positifs, alors la série $\sum_{n\geq 0} u_n$ est convergente si et seulement si elle est majorée, c'est-à-dire si et seulement si $\sum_{n\geq 0} u_n < \infty$.

Théorème 1.3.3. *(Critère de Cauchy) Si une série converge absolument, alors elle converge.*

La réciproque du résultat ci-dessus n'est pas vraie.

Proposition 1.3.4. *Soit (u_n) une suite. Si la série $\sum_{n\geq 0} u_n$ converge, alors la suite (u_n) tend vers 0.*

La réciproque de la proposition ci-dessus n'est pas vraie : par exemple, la série $\sum_{n\geq 1} \frac{1}{n}$ diverge et pourtant $\frac{1}{n} \underset{n\to\infty}{\longrightarrow} 0$.

Exemple 1.3.5.

1. Soit (u_n) la suite définie par $u_n = x^n$. La série $\sum_{n\geq 0} u_n$ est convergente si et seulement si $|x| < 1$. La somme de la série vaut

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

On parle de série géométrique.

2. Soit (u_n) la suite définie par $u_n = \frac{x^n}{n!}$. La série $\sum_{n\geq 0} u_n$ est convergente pour tout $x \in \mathbf{R}$. La somme de la série vaut

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

3. Soit (u_n) la suite définie par $u_n = \frac{1}{n^x}$. La série $\sum_{n\geq 0} u_n$ est convergente pour tout $x > 1$. La somme de la série n'a pas d'expression analytique et est connue sous le nom de fonction zêta, c'est-à-dire $\zeta(x) = \sum_{n=0}^{\infty} \frac{1}{n^x}$, pour tout $x > 1$. On parle de série de Riemann.
4. Soit (u_n) la suite définie par $u_n = \frac{(-1)^n}{n}$. La série $\sum_{n\geq 0} u_n$ est convergente mais pas absolument convergente. Il s'agit d'un cas particulier de série alternée.

L'essentiel

- calculer des intersections, réunions, produits, complémentaires, images directes et réciproques d'ensembles ;
- savoir dénombrer ;
- déterminer si un ensemble est dénombrable ou non ;
- déterminer si une série converge et, le cas échéant, calculer la somme.

Chapitre 2

Evénements

Sommaire

2.1 Tribus, événements, probabilités

Définition 2.1.1. *Soit Ω un ensemble non vide. On appelle* ***tribu*** *sur Ω toute partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ satisfaisant les quatre propriétés suivantes :*

(i) $\Omega \in \mathcal{A}$ *et* $\emptyset \in \mathcal{A}$ *;*

(ii) $\mathcal{A}$ *est stable par passage au complémentaire : pour tout $A \in \mathcal{A}$, on a $A^c \in \mathcal{A}$;*

(iii) $\mathcal{A}$ *est stable par réunion dénombrable : pour toute suite (A_n) d'éléments de $\mathcal{A}$, on a*

$$\bigcup_{n\geq 0} A_n \in \mathcal{A};$$

(iv) $\mathcal{A}$ *est stable par intersection dénombrable : pour toute suite (A_n) d'éléments de $\mathcal{A}$, on a*

$$\bigcap_{n\geq 0} A_n \in \mathcal{A}.$$

Exemple 2.1.2.

1. Soit Ω un ensemble. Les familles $\mathcal{A} = \mathcal{P}(\Omega)$ et $\mathcal{A}' = \{\emptyset, \Omega\}$ sont des tribus sur Ω.
2. Prenons $\Omega = \mathbf{R}$ et $\mathcal{A} = \{\emptyset, [0, 1], \mathbf{R} \setminus [0, 1], \mathbf{R}\}$. Alors $\mathcal{A}$ est une tribu.
3. Prenons $\Omega = \{1, 2\}$ et $\mathcal{A} = \{\emptyset, \{1\}, \{1, 2\}\}$. Alors $\mathcal{A}$ n'est pas une tribu.

Remarque 2.1.3. Pour montrer qu'une famille est une tribu, il suffit de vérifier que : *(i')* $\Omega \in \mathcal{A}$ ou $\emptyset \in \mathcal{A}$; *(ii')* $\mathcal{A}$ est stable par passage au complémentaire ; *(iii')* $\mathcal{A}$ est stable par réunion ou intersection dénombrable.

En L2, nous ne travaillerons que sur des univers Ω dénombrables. En pratique, la tribu que nous considérerons sera toujours $\mathcal{P}(\Omega)$. En L3, des univers plus généraux seront considérés et les tribus ne seront plus nécessairement $\mathcal{P}(\Omega)$.

Définition 2.1.4. *Soient Ω un ensemble et $\mathcal{A}$ une tribu sur Ω :*
- *l'ensemble Ω s'appelle l'**univers** (on l'appelle aussi l'ensemble de tous les possibles) ;*
- *le couple $(\Omega, \mathcal{A})$ s'appelle un **espace probabilisable** ;*
- *les éléments ω de Ω s'appellent des **éventualités** ;*
- *les éléments A de $\mathcal{A}$ s'appellent des **événements** ;*
- *on dit que deux événements A et B sont **incompatibles** si $A \cap B = \emptyset$;*

Exemple 2.1.5. 1. On lance un dé au hasard numéroté de 1 à 6 et on regarde la valeur obtenue. L'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}$. Les événements

$$A = \{\text{le dé est pair}\} = \{2, 4, 6\} \quad \text{et} \quad B = \{5\}$$

sont incompatibles. Le nombre 5 est une éventualité.

2. On lance deux dés (l'un après l'autre par exemple) au hasard numérotés de 1 à 6. L'univers est

$$\Omega = \{1, 2, 3, 4, 5, 6\}^2 = \{(i, j) : 1 \leq i, j \leq 6\} = \{(1, 1), (1, 2), (1, 3), \dots, (6, 5), (6, 6)\}.$$

La première coordonnée d'un couple indique la valeur du premier dé, et la seconde du second dé. Les événements

$$A = \{\text{la valeur du } 1^{er} \text{ dé est le double de celle du second}\} = \{(2, 1), (4, 2), (6, 3)\}$$

et

$$B = \{\text{la somme des valeurs obtenues par les deux dés est } 3\} = \{(1, 2), (2, 1)\}$$

ne sont pas incompatibles : l'éventualité $(2, 1)$ est commune à ces deux événements.

Définition 2.1.6. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable. On appelle (mesure de) **probabilité** sur $(\Omega, \mathcal{A})$ toute application $\mathbb{P} : \mathcal{A} \to [0, 1]$ telle que :*

(i) $\mathbb{P}(\Omega) = 1$;

(ii) pour toute suite d'événements (A_n) deux à deux disjoints (cad tels que $A_n \cap A_m = \emptyset$ pour tous $n \neq m$),

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

Exemple 2.1.7. 1. Soit Ω un ensemble fini. Prenons $\mathcal{A} = \mathcal{P}(\Omega)$. L'application $\mathbb{P}$ définie, pour tout $A \in \mathcal{P}(\Omega)$, par $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$ est une probabilité.

2. Prenons $\Omega = \mathbf{N}$, $\mathcal{A} = \mathcal{P}(\mathbf{N})$. L'application $\mathbb{P}$ définie par

$$\mathbb{P}(A) = \begin{cases} 1 & \text{si } 5 \in A \\ 0 & \text{sinon} \end{cases}$$

est une probabilité.

Définition 2.1.8. *Soit $(\Omega, \mathcal{A})$ un espace probabilisable et $\mathbb{P}$ une probabilité sur $(\Omega, \mathcal{A})$. Le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ s'appelle un **espace probabilisé**.*

Remarque 2.1.9. Supposons que Ω soit dénombrable et que $\mathcal{A} = \mathcal{P}(\Omega)$. Alors, pour tout événement $A \in \mathcal{A}$, on a

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

Exemple 2.1.10. Prenons $\Omega = \mathbf{N}$, $\mathcal{A} = \mathcal{P}(\mathbf{N})$ et $A = \{1, 2, 3\}$. Alors pour toute probabilité $\mathbb{P}$ sur $\mathbf{N}$, on a $\mathbb{P}(A) = \mathbb{P}(\{1, 2, 3\}) = \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) + \mathbb{P}(\{3\})$.

Définition 2.1.11. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On dit qu'un événement $A \in \mathcal{A}$ est :*
- ***négligeable*** *si $\mathbb{P}(A) = 0$;*
- ***presque sûr*** *si $\mathbb{P}(A) = 1$.*

Remarque 2.1.12.
1. Si $A = \emptyset$, alors A est négligeable. La réciproque n'est pas vraie.
2. Si $A = \Omega$, alors A est presque sûr. La réciproque n'est pas vraie.

Définition 2.1.13. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On appelle* ***système complet d'événements*** *toute suite (finie ou infinie) d'événements (A_n) telle que :*

(i) pour tout n, on a $\mathbb{P}(A_n) \neq 0$;

(ii) pour tous $m \neq n$, on a $\mathbb{P}(A_m \cap A_n) = 0$;

(iii) $\mathbb{P}\left(\bigcup_{n \geq 0} A_n\right) = 1$.

Exemple 2.1.14. On considère un jeu classique de 52 cartes (ARDV..., pique, coeur, carreau, trèfle). Les événements $A = \{\text{la carte est un coeur}\}$, $B = \{\text{la carte est un carreau}\}$ et $C = \{\text{la carte est de couleur noire (pique, trèfle)}\}$ forment un système complet d'événements.

Proposition 2.1.15. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.*

(i) Pour tout $A \in \mathcal{A}$, on a $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(ii) Pour tous $A, B \in \mathcal{A}$ tels que $A \subset B$, on a $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$ et $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(iii) Pour tous $A, B \in \mathcal{A}$, on a

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

(iv) Pour tous $A_1, \ldots, A_n \in \mathcal{A}$ deux à deux disjoints, on a

$$\mathbb{P}\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}(A_k).$$

(v) Pour tous $A_1, \ldots, A_n \in \mathcal{A}$, on a

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leq \sum_{k=1}^{n} \mathbb{P}(A_k).$$

(vi) Pour toute suite d'événements (A_n), on a

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Proposition 2.1.16. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et (A_n) une suite d'événements.*

(i) Si (A_n) est croissante pour l'inclusion, cad si $A_n \subset A_{n+1}$ pour tout n, alors la suite $(\mathbb{P}(A_n))_{n\geq 1}$ est croissante et sa limite est :

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right).$$

(ii) Si (A_n) est décroissante pour l'inclusion, cad si $A_n \supset A_{n+1}$ pour tout n, alors la suite $(\mathbb{P}(A_n))_{n\geq 1}$ est décroissante et sa limite est :

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right).$$

2.2 Evénements indépendants

Définition 2.2.1. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On dit que deux événements $A, B \in \mathcal{A}$ sont **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Définition 2.2.2. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_i)_{i\in I}$ une famille d'événements. On dit que les événements $(A_i)_{i\in I}$ sont :*

- ***deux à deux indépendants** si A_i et A_j sont indépendants pour tous $i, j \in I$, $i \neq j$;*
- ***mutuellement indépendants** si pour toute famille finie $J \subset I$,*

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

Exemple 2.2.3. On lance deux pièces au hasard et on note P (resp. F) lorsqu'on on obtient pile (resp. face). L'univers est $\Omega = \{(P,P),(P,F),(F,P),(F,F)\}$. On considère les événements suivants :

$$A = \{\text{le premier lancer est pile}\} = \{(P,P),(P,F)\};$$

$$B = \{\text{le second lancer est face}\} = \{(P,F),(F,F)\};$$

$$C = \{\text{le résultat est le même aux deux lancers}\} = \{(P,P),(F,F)\}.$$

On a $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$ et $\mathbb{P}(A\cap B) = \mathbb{P}(A\cap C) = \mathbb{P}(B\cap C) = \frac{1}{4}$. Les événements A, B, C sont donc deux à deux indépendants. En revanche, ils ne sont pas mutuellement indépendants puisque $\mathbb{P}(A\cap B\cap C) = 0 \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Remarque 2.2.4.

1. Il ne faut pas confondre les notions d'indépendance (A et B sont indépendants si $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$) et d'incompatibilité ($A$ et B sont incompatibles si $A \cap B = \emptyset$) de deux événements.
2. La notion d'indépendance est relative à la probabilité : deux événements peuvent être indépendants relativement à une probabilité $\mathbb{P}$ mais pas à une probabilité $\mathbb{P}'$.
3. Si $(A_i)_{i\in I}$ est une famille d'événements mutuellement indépendants (resp. deux à deux indépendants) alors $(A_i^c)_{i\in I}$ est une famille d'événements mutuellement indépendants (resp. deux à deux indépendants).

2.3 Probabilités conditionnelles

Proposition 2.3.1. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* $A \in \mathcal{A}$ *un événement non négligeable. L'application* $\mathbb{P}_A : \mathcal{A} \to [0,1]$ *définie, pour tout* $B \in \mathcal{A}$*, par :*

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

est une probabilité.

Définition 2.3.2. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé, et* A *et* B *deux événements tels que* A *n'est pas négligeable. On appelle probabilité de* B ***conditionnellement*** *à* A *(ou probabilité de* B *sachant* A*) la quantité définie par :*

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

Proposition 2.3.3. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé, et* A *et* B *deux événements tels que* A *n'est pas négligeable. Alors* A *et* B *sont indépendants si et seulement si* $\mathbb{P}(B|A) = \mathbb{P}(B)$*.*

Proposition 2.3.4. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* $(A_k)_{k \leq n}$ *une famille d'intersection non négligeable. Alors*

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_k\right) = \mathbb{P}\left(A_n \middle| \bigcap_{k=1}^{n-1} A_k\right) \mathbb{P}\left(A_{n-1} \middle| \bigcap_{k=1}^{n-2} A_k\right) \cdots \mathbb{P}(A_2|A_1)\,\mathbb{P}(A_1).$$

Théorème 2.3.5. *(Formule des probabilités totales) Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* (A_n) *un système complet d'événements non négligeables. Alors, pour tout* $B \in \mathcal{A}$*,*

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B \cap A_n) = \sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\,\mathbb{P}(A_n).$$

Théorème 2.3.6. *(Formule de Bayes) Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé. Soit* $B \in \mathcal{A}$ *un événement non négligeable.*

(i) Pour tout événement $A \in \mathcal{A}$ *non négligeable, on a*

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

(ii) Pour tout système complet d'événements (A_n) *non négligeables, on a*

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\,\mathbb{P}(A_j)}{\sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\,\mathbb{P}(A_n)}.$$

Exemple 2.3.7. Deux opérateurs de saisie, A et B, entrent respectivement 100 et 200 tableaux sur informatique. Les tableaux de A comportent des fautes dans 5.2% des cas et ceux de B dans 6,7% des cas. On prend un tableau au hasard. Il comporte des fautes. On souhaite calculer la probabilité pour que A se soit occupé de ce tableau. Pour cela, on considère les événements :

- T_A : "le tableau est entré par A" ;
- $T_B = T_A^c$: "le tableau est entré par B" ;
- F : "le tableau comporte des fautes".

D'après la formule de Bayes, on a

$$\begin{aligned}\mathbb{P}\left(T_A|F\right) &= \frac{\mathbb{P}\left(F|T_A\right)\mathbb{P}\left(T_A\right)}{\mathbb{P}\left(F|T_A\right)\mathbb{P}\left(T_A\right)+\mathbb{P}\left(F|T_B\right)\mathbb{P}\left(T_B\right)}\\ &= \frac{0.052\times 1/3}{0.052\times 1/3+0.067\times 2/3}\\ &= 0.279.\end{aligned}$$

La théorie des probabilités s'étend au cas d'un univers Ω non dénombrable. La théorie de la mesure (abordée en L3 et M1) donne des outils dans un cadre général et permet de généraliser la notion de probabilité. Un des premiers résultats est de construire une tribu $\mathcal{B}(\mathbf{R})$ (appelée tribu borélienne) ainsi qu'une mesure sur $\mathbf{R}$ (appelée mesure de Lebesgue).

L'essentiel

- avoir en tête le vocabulaire probabiliste (univers, événements, probabilité, probabilité conditionnelle, indépendance) ;
- calculer des probabilités et des probabilités conditionnelles ;
- étudier l'indépendance d'une famille d'événements.

Chapitre 3

Variables aléatoires

Sommaire

3.1 Variable aléatoire, loi

Définition 3.1.1. *Soient* $(\Omega, \mathcal{A})$ *et* $(E, \mathcal{E})$ *deux espaces probabilisables. On appelle* ***variable aléatoire discrète*** *toute fonction* $X : \Omega \to E$ *telle que* $X(\Omega)$ *soit dénombrable et telle que :*

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

L'ensemble $X(\Omega)$ s'appelle l'**espace d'état** ou le **support** de X. Dans tout le chapitre, on suppose que pour tout $x \in E$, on a $\{x\} \in \mathcal{E}$ et $X^{-1}(\{x\}) \in \mathcal{A}$. En pratique, de telles conditions sont satisfaites en L2 car les espaces Ω et E sont dénombrables et les tribus sous-jacentes sont $\mathcal{A} = \mathcal{P}(\Omega)$ et $\mathcal{E} = \mathcal{P}(E)$.

Exemple 3.1.2. On lance une pièce au hasard. Si l'on tombe sur pile (P), on gagne 1 euro, et si c'est sur face (F) on en gagne 0. Ici, l'univers est $\Omega = \{P, F\}$ et la variable aléatoire qui nous intéresse est :

$$\begin{aligned} X :&\{P, F\} \to \{0, 1\} \\ &P \mapsto 1 \\ &F \mapsto 0. \end{aligned}$$

Proposition 3.1.3. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé. L'ensemble des variables aléatoires discrètes définies sur* Ω *à valeurs dans* $\mathbf{R}$ *est un espace vectoriel.*

Proposition 3.1.4. *Soient* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* $(E, \mathcal{E})$ *un espace probabilisable. Soit* X *une variable aléatoire discrète à valeurs dans* E*. Soit* $\mathbb{P}_X : \mathcal{E} \to [0, 1]$ *l'application définie, pour tout* $A \in \mathcal{E}$*, par :*

$$\mathbb{P}_X(A) = \mathbb{P}\left(X^{-1}(A)\right).$$

Alors l'application $\mathbb{P}_X$ *est une probabilité sur l'espace probabilisable* $(E, \mathcal{E})$*.*

Par définition de $\mathbb{P}_X$, on a donc $\mathbb{P}_X(A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$. Il est d'usage, en probabilités, d'alléger les écritures en prenant la notation suivante :

$$\{X \in A\} = \{\omega \in \Omega : X(\omega) \in A\} = X^{-1}(A).$$

Par exemple, si X est à valeurs réelles, l'événement $\{X \leq x\}$ est égal à $\{\omega \in \Omega : X(\omega) \leq x\}$. On écrit également $\mathbb{P}(X \in A)$ au lieu de $\mathbb{P}(\{X \in A\})$ en omettant les parenthèses. En d'autres termes, on a

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}\left(X^{-1}(A)\right) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}) \\ &= \mathbb{P}(\{X \in A\}) \\ &= \mathbb{P}(X \in A). \end{aligned}$$

Définition 3.1.5. *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* X *une variable aléatoire discrète à valeurs dans un ensemble* E*. L'application* $\mathbb{P}_X$ *définie ci-dessus s'appelle la* ***loi*** *de* X*.*

La loi d'une variable aléatoire est donc, par définition, une probabilité sur l'ensemble d'arrivée de la variable aléatoire.

Proposition 3.1.6. *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et* X *une variable aléatoire discrète à valeurs dans un ensemble* E*. La loi de* X *est caractérisée de manière unique par les valeurs* $\mathbb{P}(X = x)$ *pour tout* $x \in X(\Omega)$*. Plus précisément, pour tout* $A \in \mathcal{E}$*, on a*

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

En particulier, on a $\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = 1$.

Exemple 3.1.7. On jette deux dés parfaits au hasard, numérotés de 1 à 6, et on s'intéresse à la somme des valeurs obtenues par les deux dés. L'univers est $\Omega = \{1, \ldots, 6\}^2$, la tribu est $\mathcal{A} = \mathcal{P}(\Omega)$ et la variable aléatoire qui nous intéresse est :

$$\begin{aligned} X : &\{1, \ldots, 6\}^2 \longrightarrow \{2, \ldots, 12\} \\ &(\omega_1, \omega_2) \longmapsto \omega_1 + \omega_2. \end{aligned}$$

Sa loi $\mathbb{P}_X$ est une probabilité sur $X(\Omega) = \{2, \ldots, 12\}$. On a par exemple :

$$\mathbb{P}_X(\{2\}) = \mathbb{P}(X = 2) = \mathbb{P}\left(\text{le } 1^{er} \text{ dé donne 1, le } 2^{nd} \text{ donne 1}\right) = \frac{1}{36};$$

$$\begin{aligned} \mathbb{P}_X(\{3\}) &= \mathbb{P}(X = 3) \\ &= \mathbb{P}\left(\text{le } 1^{er} \text{ dé donne 1, le } 2^{nd} \text{ donne 2}\right) + \mathbb{P}\left(\text{le } 1^{er} \text{ donne 2, le } 2^{nd} \text{ dé donne 1}\right) \\ &= \frac{2}{36}; \end{aligned}$$

Plus généralement, on a les résultats suivants :

Définition 3.1.8. *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace probabilisé et soient* $(E_1, \mathcal{E}_1)$*, ...,* $(E_p, \mathcal{E}_p)$ *des espaces probabilisables.*

sommes possibles	2	3	4	5	6	7	8	9	10	11	12
probabilités	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

*(i) Soient $X_1, \dots, X_p$ des variables aléatoires discrètes, définies sur Ω, à valeurs dans $E_1, \dots, E_p$ respectivement. On appelle **loi jointe** de $X_1, \dots, X_p$ la loi du p-uplet $X = (X_1, \dots, X_p)$. On la note $\mathbb{P}_X$.*

*(ii) Soit $X = (X_1, \dots, X_p)$ une variable aléatoire, définie sur Ω, à valeurs dans $E_1 \times \dots \times E_p$. Pour tout $1 \leq k \leq p$, on appelle **k-ième marginale** de X la loi de la variable aléatoire X_k.*

Proposition 3.1.9. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et soient $(E_1, \mathcal{E}_1)$, ..., $(E_p, \mathcal{E}_p)$ des espaces probabilisables. Soit $X = (X_1, \dots, X_p)$ une variable aléatoire à valeurs dans $E_1 \times \dots \times E_p$ de loi $\mathbb{P}_X$. Alors, pour tout $1 \leq i \leq p$, la loi de X_i est donnée par :*

$$\forall A \in \mathcal{E}_i, \quad \mathbb{P}_{X_i}(A) = \mathbb{P}_X(E_1 \times E_2 \times \dots \times E_{i-1} \times A \times E_{i+1} \times \dots \times E_p).$$

Le résultat ci-dessus implique qu'on peut avoir les lois marginales à partir de la loi jointe. La réciproque est en revanche fausse. En d'autres termes, il est possible que l'on ait deux couples de variables aléatoires (X_1, X_2) et (Y_1, Y_2) telles que $\mathbb{P}_{X_1} = \mathbb{P}_{Y_1}$ et $\mathbb{P}_{X_2} = \mathbb{P}_{Y_2}$ mais telles que $\mathbb{P}_{(X_1,X_2)} \neq \mathbb{P}_{(Y_1,Y_2)}$.

Définition 3.1.10. *Une variable aléatoire X est dite **déterministe** si, presque sûrement, elle est égale à une certaine valeur ; en d'autres termes il existe c tel que $\mathbb{P}(X = c) = 1$.*

3.2 Lois usuelles

Définition 3.2.1. *Soient E un ensemble fini et $X : \Omega \to E$ une variable aléatoire discrète. On dit que X suit la **loi uniforme** sur E si :*

$$\forall x \in E, \quad \mathbb{P}(X = x) = \frac{1}{|E|}.$$

On note alors $X \sim \mathcal{U}(E)$.

La loi uniforme apparait naturellement comme la modélisation d'un choix d'un élément au hasard dans un ensemble fini en situation d'équiprobabilité. Par exemple, on prend une carte au hasard dans un jeu (classique) de 52 cartes. On 1/52 chance de tirer l'as de pique ; 1/52 chance de tirer le 3 de trèfle ; 1/52 de chance de tirer le valet de coeur.

Définition 3.2.2. *Soient $p \in [0, 1]$ et $X : \Omega \to \{0, 1\}$ une variable aléatoire discrète. On dit que X suit la **loi de Bernoulli** de paramètre p si :*

$$\mathbb{P}(X = 1) = p \quad \textit{et} \quad \mathbb{P}(X = 0) = 1 - p.$$

On note alors $X \sim \mathcal{B}(p)$.

La loi de Bernoulli modélise le succès ou l'échec d'une expérience.

Exemple 3.2.3. On lance une pièce au hasard : si l'on tombe sur pile (avec probabilité p), on considère que l'on a un succès et on affecte la valeur 1 ; et dans le cas opposé (cad si on tombe sur face), on considère qu'on a un échec et on affecte la valeur 0. La variable aléatoire considérée est :

$$\begin{aligned} X :&\{P,F\} \to \{0,1\} \\ &P \mapsto 1 \\ &F \mapsto 0. \end{aligned}$$

On a $X \sim \mathcal{B}(p)$.

Définition 3.2.4. *Soient $p \in [0,1]$, $n \geq 1$ et $X : \Omega \to \{0,1,\ldots,n\}$ une variable aléatoire discrète. On dit que X suit la **loi binomiale** de paramètres n, p si :*

$$\forall 0 \leq k \leq n, \quad \mathbb{P}\left(X = k\right) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note alors $X \sim \mathcal{B}(n,p)$.

La loi binomiale intervient dans le comptage du nombre de succès d'une expérience aléatoire. Plus précisément, la somme de n variables aléatoires indépendantes (on définira cette notion dans la section suivante) de loi de Bernoulli de paramètre p est une variable aléatoire de loi binomiale de paramètres n et p.

Exemple 3.2.5. On lance 10 fois une pièce (indépendamment des autres lancers) qui a une probabilité p de tomber sur pile. L'univers est $\Omega = \{P,F\}^{10}$. On note X la variable aléatoire X qui désigne le nombre de fois que l'on tombe sur pile. En d'autres termes

$$\begin{aligned} X :&\{P,F\}^{10} \longrightarrow \{0,\ldots,10\} \\ &(\omega_1,\ldots,\omega_{10}) \longmapsto \#\{1 \leq i \leq 10 : \omega_i = P\}. \end{aligned}$$

On a $X \sim \mathcal{B}(10,p)$.

Remarque 3.2.6. La loi binomiale intervient naturellement dans des problèmes de tirages de boules *avec remise*. Considérons, par exemple, la situation suivante : on a N boules dans une urne, N_B d'entre elles sont blanches et $N_R = N - N_B$ sont rouges.

- On prélève au hasard, et *avec remise*, $n \leq N$ boules dans l'urne. On désigne par X le nombre de boules blanches que l'on tire. Alors X suit la loi $\mathcal{B}(n,p)$, avec $p = \frac{N_B}{N}$, cad pour tout $k \leq n$,

  $$\mathbb{P}\left(X = k\right) = \binom{n}{k} p^k (1-p)^{n-k}.$$

- On prélève au hasard, et *sans remise*, $n \leq N$ boules dans l'urne. On note Y le nombre de boules blanches que l'on tire. Alors la loi de Y est donnée par :

  $$\mathbb{P}\left(Y = k\right) = \frac{\binom{N_B}{k}\binom{N_R}{n-k}}{\binom{N}{n}},$$

  pour tout $k \leq n$. On dit que Y suit une loi hypergéométrique.

Définition 3.2.7. *Soit $p \in [0,1]$ et $X : \Omega \to \mathbf{N}^*$ une variable aléatoire discrète. On dit que X suit la **loi géométrique** de paramètre p si :*

$$\forall k \in \mathbf{N}^*, \quad \mathbb{P}\left(X = k\right) = p(1-p)^{k-1}.$$

On note alors $X \sim \mathcal{G}(p)$.

La loi géométrique apparait comme l'instant du premier succès lorsqu'on itère indépendamment une même expérience aléatoire ayant une probabilité p de succès et $1-p$ d'échec.

Exemple 3.2.8. On lance à plusieurs reprises une pièce qui a une probabilité $p \in]0,1[$ de tomber sur pile. On désigne par X le numéro du lancer où on a obtenu pour la première fois pile. On a $X \sim \mathcal{G}(p)$.

Définition 3.2.9. *Soit $\lambda > 0$ et $X : \Omega \to \mathbf{N}$ une variable aléatoire discrète. On dit que X suit la **loi de Poisson** de paramètre λ si :*

$$\forall k \in \mathbf{N}, \quad \mathbb{P}(X = k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

On note alors $X \sim \mathcal{P}(\lambda)$.

La loi de Poisson est souvent utilisée pour modéliser le nombre d'occurrences d'un phénomène aléatoire extrême/rare. Le résultat suivant garantit qu'une loi binomiale de paramètres n, p_n, avec $np_n \to \lambda > 0$, est proche d'une loi de Poisson de paramètre λ.

Proposition 3.2.10. *Soit $\lambda > 0$ et soit (p_n) une suite de nombres réels compris dans $[0,1]$ telle que $np_n \underset{n\to\infty}{\longrightarrow} \lambda$. Alors*

$$\forall k \in \mathbf{N}, \quad \binom{n}{k} p_n^k (1-p_n)^{n-k} \underset{n\to\infty}{\longrightarrow} e^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

En d'autres termes, si X est une variable aléatoire telle que $X \sim \mathcal{P}(\lambda)$ et si (X_n) est une suite de variables aléatoires telles que $X_n \sim \mathcal{B}(n, p_n)$ pour tout $n \geq 1$, alors

$$\forall k \in \mathbf{N}, \quad \mathbb{P}(X_n = k) \underset{n\to\infty}{\longrightarrow} \mathbb{P}(X = k).$$

On dit qu'une suite de variables aléatoires (X_n) qui satisfait l'équation ci-dessus converge en loi vers la variable aléatoire X. On abordera ce type de notion dans le dernier chapitre.

3.3 Indépendance de variables aléatoires

Définition 3.3.1. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. On dit qu'une famille de variables aléatoires discrètes $(X_i)_{i\in I}$ est **indépendante** si, pour toute partie finie $J \subset I$, et toute famille $(A_j)_{j\in J}$ avec $A_j \subset X_j(\Omega)$ pour tout $j \in J$, on a*

$$\mathbb{P}(\forall j \in J,\, X_j \in A_j) = \prod_{j\in J} \mathbb{P}(X_j \in A_j).$$

Proposition 3.3.2. *Soitent X et Y deux variables aléatoires discrètes, définies sur Ω. Les propriétés suivantes sont équivalentes :*

(i) X et Y sont indépendantes ;

(ii) $\forall x \in X(\Omega), \forall y \in Y(\Omega)$, $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\,\mathbb{P}(Y = y)$.

Proposition 3.3.3. *Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, I un ensemble et $(X_i)_{i\in I}$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$. On suppose que les variables aléatoires sont indépendantes.*

(i) Pour tout $i \in I$, désignons par f_i une application définie sur $X_i(\Omega)$. Alors les variables aléatoires $(f_i(X_i))_{i\in I}$ sont indépendantes.

(ii) Soient $J, J' \subset I$ deux sous-ensembles disjoints de I. Alors les deux vecteurs $(X_i)_{i\in J}$ et $(X_j)_{j\in J'}$ sont indépendants.

Définition 3.3.4. *Soient X, Y deux variables aléatoires discrètes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans E et F respectivement et $B \subset F$ tel que $\mathbb{P}(Y \in B) > 0$. La **loi conditionnelle** de X sachant que $Y \in B$ est la probabilité conditionnelle $\mathbb{P}_X^{\{Y\in B\}}$ sur $(E, \mathcal{P}(E))$ définie, pour tout $A \in \mathcal{P}(E)$, par :*

$$\mathbb{P}_X^{\{Y\in B\}}(A) = \frac{\mathbb{P}(X \in A, Y \in B)}{\mathbb{P}(Y \in B)}.$$

On note plus simplement $\mathbb{P}(X \in A|Y \in B) = \mathbb{P}_X^{\{Y\in B\}}(A)$. Etant donné que les variables aléatoires X et Y sont discrètes, la loi conditionnelle de X sachant que $Y \in B$ est caractérisée par les probabilités

$$\mathbb{P}(X = x|Y \in B) = \frac{\mathbb{P}(X = x, Y \in B)}{\mathbb{P}(Y \in B)}$$

pour tout $x \in X(\Omega)$.

Le théorème suivant, admis, garantit l'existence de familles de variables aléatoires indépendantes (la preuve repose sur l'existence d'une mesure produit, telle qu'on l'aborde en M1).

Théorème 3.3.5. *Soit $(E_i)_{i\in I}$ une famille d'ensembles dénombrables et $(\mathbb{P}_i)_{i\in I}$ une famille de probabilités sur ces ensembles. Alors il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et des variables aléatoires $X_i : (\Omega, \mathcal{A}) \to E_i$ telles que :*

- *pour tout $i \in I$, la loi de X_i est $\mathbb{P}_i$;*
- *les variables aléatoires $(X_i)_{i\in I}$ sont indépendantes.*

L'essentiel

- calculer la loi d'une variable aléatoire ;
- reconnaitre les lois usuelles (uniforme, Bernoulli, binomiale, Poisson, géométrique) ;
- étudier l'indépendance d'une famille de variables aléatoires.

Chapitre 4

Espérance de variables aléatoires réelles

Sommaire

Dans ce chapitre, nous nous limitons aux variables aléatoires discrètes réelles (cad à valeurs dans un sous-ensemble dénombrable de $\mathbf{R}$).

4.1 Espérance

Définition 4.1.1. *Soit X une variable aléatoire discrète à valeurs dans $\mathbf{R}$.*

- *On dit que X admet une espérance si la famille $(x\mathbb{P}(X=x))_{x\in X(\Omega)}$ est sommable, cad si $\sum_{x\in X(\Omega)} |x|\mathbb{P}(X=x) < \infty$.*
- *Lorsque X admet une espérance, on appelle **espérance** de X le nombre réel défini par :*

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

Une variable aléatoire dont l'espérance est nulle est dite **centrée**. Dans le cas d'une variable positive, on peut toujours donner un sens à la notation $\mathbb{E}[X]$, même si la famille $(x\mathbb{P}(X=x))_{x\in X(\Omega)}$ n'est pas sommable, en convenant que $\mathbb{E}[X] = +\infty$ dans ce cas.

Remarque 4.1.2. Deux variables aléatoires discrètes (réelles) qui ont la même loi ont la même espérance. La réciproque n'est pas vraie.

Exemple 4.1.3.

1. Soit $X \sim \mathcal{U}(\{1,2,3,4\})$. Alors $\mathbb{E}[X] = 2.5$.
2. Soit $X \sim \mathcal{B}(p)$, $p \in [0,1]$. Alors $\mathbb{E}[X] = p$.
3. Soit $X \sim \mathcal{B}(n,p)$, $n \geq 1$, $p \in [0,1]$. Alors $\mathbb{E}[X] = np$.
4. Soit $X \sim \mathcal{G}(p)$, $p \in]0,1]$. Alors $\mathbb{E}[X] = \frac{1}{p}$.
5. Soit $X \sim \mathcal{P}(\lambda)$, $\lambda > 0$. Alors $\mathbb{E}[X] = \lambda$.

Remarque 4.1.4. Soit $X : (\Omega, \mathcal{A}) \to E$ une variable aléatoire discrète et $A \subset E$. Alors $\mathbb{E}[\mathbf{1}_{X \in A}] = \mathbb{P}(X \in A)$.

Théorème 4.1.5. *(Théorème de transfert) Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, E un ensemble, $X : (\Omega, \mathcal{A}) \to E$ une variable aléatoire discrète et $f : X(\Omega) \to \mathbf{R}$ une application. Si la famille $(f(x)\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable (cad si $\sum_{x \in X(\Omega)} |f(x)|\mathbb{P}(X = x) < \infty$), alors $f(X)$ admet une espérance et celle-ci est donnée par :*

$$\mathbb{E}[f(X)] = \sum_{x \in X(\Omega)} f(x)\mathbb{P}(X = x).$$

Proposition 4.1.6. *Soient X, Y deux variables aléatoires discrètes réelles qui admettent une espérance. Alors*

(i) linéarité : pour tout $\lambda \in \mathbf{R}$, la variable aléatoire $X + \lambda Y$ admet une espérance et celle-ci est donnée par :

$$\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda \mathbb{E}[Y].$$

(ii) croissance : si $X \leq Y$ presque sûrement, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

(iii) inégalité triangulaire : $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

Corollaire 4.1.7. Soient X, Y deux variables aléatoires discrètes réelles telles que $|X| \leq Y$ presque sûrement. Si Y admet une espérance, alors X admet une espérance.

Proposition 4.1.8. *Soit X une variable aléatoire discrète. Si $X \geq 0$ presque sûrement et si $\mathbb{E}[X] = 0$ alors $X = 0$ presque sûrement.*

Proposition 4.1.9. *Soient X et Y deux variables aléatoires discrètes possédant une espérance. Si X et Y sont indépendantes, alors XY admet une espérance et $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.*

La proposition suivante est une généralisation du résultat ci-dessus.

Proposition 4.1.10. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Pour tout $i \leq n$, soient $X_i : (\Omega, \mathcal{A}) \to E_i$ une variable aléatoire et $f_i : E_i \to \mathbf{R}$ une application telle que $f_i(X_i)$ admette une espérance. Supposons que les variables aléatoires $X_1, \ldots, X_n$ soient indépendantes. Alors $\prod_{i=1}^n f_i(X_i)$ admet une espérance et*

$$\mathbb{E}\left[\prod_{i=1}^n f_i(X_i)\right] = \prod_{i=1}^n \mathbb{E}[f_i(X_i)].$$

4.2 Variance et moments d'ordre supérieur

Définition 4.2.1. *Soit X une variable aléatoire discrète réelle. On dit que X admet un moment d'ordre $k \geq 1$ si la variable aléatoire X^k admet une espérance. Le **moment d'ordre** k est alors le réel $\mathbb{E}\left[X^k\right]$.*

Définition 4.2.2. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2 cad telle que $\mathbb{E}\left[X^2\right] < \infty$. On appelle :*

- ***variance** de X le nombre réel positif*

$$\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right];$$

- ***écart-type*** *de X le nombre réel positif*

$$\sigma(X) = \sqrt{\mathbb{V}[X]}.$$

Une variable aléatoire centrée dont la variance est égale à 1 est dite **réduite**.

Exemple 4.2.3. 1. Soit $X \sim \mathcal{B}(p)$, $p \in [0,1]$. Alors $\mathbb{V}[X] = p(1-p)$.

2. Soit $X \sim \mathcal{B}(n,p)$, $n \geq 1$, $p \in [0,1]$. Alors $\mathbb{V}[X] = np(1-p)$.
3. Soit $X \sim \mathcal{G}(p)$, $p \in]0,1]$. Alors $\mathbb{V}[X] = \frac{1-p}{p^2}$.
4. Soit $X \sim \mathcal{P}(\lambda)$, $\lambda > 0$. Alors $\mathbb{V}[X] = \lambda$.

Proposition 4.2.4. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors X admet une espérance.*

Proposition 4.2.5. *(Formule de König-Huygens) Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors*

$$\mathbb{V}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2.$$

Proposition 4.2.6. *Soit X une variable aléatoire discrète admettant un moment d'ordre 2. Alors*

(i) $\mathbb{V}[X] = 0$ *si et seulement si X est déterministe.*

(ii) Soient $a, b \in \mathbf{R}$. Alors

$$\mathbb{V}[aX + b] = a^2\,\mathbb{V}[X].$$

Proposition 4.2.7. *Soient X et Y deux variables aléatoires discrètes réelles indépendantes admettant des moments d'ordre 2. Alors* $\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y]$.

L'hypothèse d'indépendance dans la proposition ci-dessus est essentielle. Sans supposer cela, il n'est en général pas vrai que $\mathbb{V}[X+Y] = \mathbb{V}[X] + \mathbb{V}[Y]$. Par exemple, si X est une variable aléatoire réelle et $Y = 2X$, alors

$$\mathbb{V}[X+Y] = \mathbb{V}[3X] = 9\,\mathbb{V}[X]$$

mais

$$\mathbb{V}[X] + \mathbb{V}[Y] = \mathbb{V}[X] + \mathbb{V}[2X] = \mathbb{V}[X] + 4\,\mathbb{V}[X] = 5\,\mathbb{V}[X].$$

Définition 4.2.8. *Soient X et Y deux variables aléatoires discrètes réelles admettant des moments d'ordre 2. On appelle* ***covariance*** *de X et de Y le nombre réel*

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

Remarque 4.2.9. 1. La covariance "généralise" la notion de variance : $\mathrm{Cov}(X,X) = \mathbb{V}[X]$.

2. La covariance est bilinéaire cad $\mathrm{Cov}(X, aY + Z) = a\mathrm{Cov}(X,Y) + \mathrm{Cov}(X,Z)$.
3. La covariance est symétrique cad $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.

Définition 4.2.10. *Deux variables aléatoires dont la covariance existe et est égale à 0 sont dites* ***non corrélées.***

Proposition 4.2.11. *Si deux variables aléatoires sont indépendantes alors elles sont non corrélées.*

La réciproque de la proposition précédente n'est pas vraie. Voici un contre-exemple.

Contre-exemple 4.2.12. Soit X une variable aléatoire telle que $\mathbb{P}(X=0) = \mathbb{P}(X=1) = \mathbb{P}(X=-1) = \frac{1}{3}$. Soit Y la variable aléatoire définie par :

$$Y = \begin{cases} 0 & \text{si } X \neq 0 \\ 1 & \text{si } X = 0. \end{cases}$$

On a $\mathrm{Cov}(X,Y) = 0$ et pourtant les variables aléatoires X et Y ne sont pas indépendantes puisque

$$\mathbb{P}(X=0, Y=0) = 0 \neq \frac{2}{9} = \mathbb{P}(X=0)\,\mathbb{P}(Y=0).$$

4.3 Fonction génératrice

Définition 4.3.1. *Soit X une variable aléatoire à valeurs dans $\mathbf{N}$. On appelle* ***fonction génératrice*** *de X la somme de la série entière de coefficients $\mathbb{P}(X=k)$, cad*

$$G_X : t \mapsto \sum_{k=0}^{\infty} \mathbb{P}(X=k)\, t^k,$$

où t appartient au disque de convergence de la série entière $\sum_{k\geq 0} \mathbb{P}(X=k)\, t^k$.

Exemple 4.3.2.
1. Soit $X \sim \mathcal{B}(p)$, $p \in [0,1]$. Alors $G_X(t) = 1 - p + pt$, $t \in \mathbf{R}$.
2. Soit $X \sim \mathcal{B}(n,p)$, $n \geq 1$, $p \in [0,1]$. Alors $G_X(t) = (1-p+pt)^n$, $t \in \mathbf{R}$.
3. Soit $X \sim \mathcal{G}(p)$, $p \in]0,1]$. Alors $G_X(t) = \frac{pt}{1-(1-p)t}$, $|t| < \frac{1}{1-p}$.
4. Soit $X \sim \mathcal{P}(\lambda)$, $\lambda > 0$. Alors $G_X(t) = e^{\lambda(t-1)}$, $t \in \mathbf{R}$.

Comme $\sum_{k=0}^{\infty} \mathbb{P}(X=k) = 1$, on a $G_X(1) = 1$. En particulier, la série définissant $G_X(t)$ converge pour tout $|t| \leq 1$. Le rayon de convergence de la série entière associée à G_X vaut donc au moins 1. Il en résulte que G_X est de classe $\mathcal{C}^\infty$ sur $]-1,1[$ et que G_X est continue en -1 et en 1. On obtient également que $\mathbb{P}(X=0) = G_X(0)$ et plus généralement que

$$\mathbb{P}(X=n) = \frac{G_X^{(n)}(0)}{n!}.$$

En particulier, la série génératrice caractérise la loi. Plus précisément, on a le théorème suivant.

Théorème 4.3.3. *Soient X et Y deux variables aléatoires à valeurs dans $\mathbf{N}$ telles qu'il existe $r > 0$ vérifiant :*

$$\forall t \in [0,r], \quad G_X(t) = G_Y(t).$$

Alors X et Y ont la même loi.

Proposition 4.3.4. *Soit X une variable aléatoire à valeurs dans $\mathbf{N}$. Pour tout $t \in [-1,1]$, on a*

$$G_X(t) = \mathbb{E}\left[t^X\right].$$

Proposition 4.3.5. *Soient X une variable aléatoire à valeurs dans $\mathbf{N}$ et $k \geq 1$. La variable aléatoire X admet un moment d'ordre k si et seulement si G_X est k-fois dérivable en 1. Dans ce cas,*

$$G_X^{(k)}(1) = \mathbb{E}\left[X(X-1)\cdots(X-(k-1))\right].$$

Proposition 4.3.6. *Soient X et Y deux variables aléatoires à valeurs dans $\mathbf{N}$ indépendantes. Alors*

$$G_{X+Y} = G_X \times G_Y.$$

4.4 Inégalités classiques

Proposition 4.4.1. *(Inégalité de Markov) Soit X une variable aléatoire réelle positive admettant une espérance. Alors, pour tout $a > 0$,*

$$\mathbb{P}\left(X \geq a\right) \leq \frac{\mathbb{E}\left[X\right]}{a}.$$

Proposition 4.4.2. *(Inégalité de Tchebychev) Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors, pour tout $a > 0$,*

$$\mathbb{P}\left(|X - \mathbb{E}\left[X\right]| \geq a\right) \leq \frac{\mathbb{V}\left[X\right]}{a^2}.$$

La proposition ci-dessus entraine que, plus la variance est petite, plus la variable aléatoire est concentrée autour de sa moyenne : en d'autres termes, elle varie peu autour de sa moyenne (d'où le nom de variance).

Proposition 4.4.3. *(Inégalité de Cauchy-Schwarz) Soient X et Y deux variables aléatoires réelles admettant un moment d'ordre 2. Alors XY admet une espérance et*

$$\mathbb{E}\left[|XY|\right] \leq \sqrt{\mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]}.$$

Le résultat ci-dessus permet notamment de montrer que le coefficient de corrélation de deux variables aléatoires X et Y admettant des moments d'ordre 2, défini par

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

vérifie l'inégalité $|\rho(X,Y)| \leq 1$.

Proposition 4.4.4. *(Inégalité de Jensen) Soit X une variable aléatoire réelle à valeurs dans l'intervalle I, admettant une espérance. Soit $\varphi : I \to \mathbf{R}$ une fonction convexe. On suppose que la variable aléatoire $\varphi(X)$ admet une espérance. Alors*

$$\varphi(\mathbb{E}\left[X\right]) \leq \mathbb{E}\left[\varphi(X)\right].$$

L'inégalité ci-dessus permet de retrouver le fait que $|\mathbb{E}\left[X\right]| \leq \mathbb{E}\left[|X|\right]$ (inégalité triangulaire) et que $(\mathbb{E}\left[X\right])^2 \leq \mathbb{E}\left[X^2\right]$.

L'essentiel

- montrer qu'une variable aléatoire possède une espérance ;
- calculer une espérance, variance, covariance ;
- connaitre les espérances, variances des lois classiques (Bernoulli, binomiale, Poisson) ;
- calculer une fonction génératrice.

Chapitre 5

Théorèmes limites

Sommaire

5.1 Différents modes de convergence

On distingue deux types de convergences : celles de type "spatial" où l'univers Ω joue un rôle essentiel, et la convergence "en loi" où seules les lois des variables aléatoires importent.

Convergences spatiales

Définition 5.1.1. *Soient (X_n) et X des variables aléatoires réelles définies sur le même espace de probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que :*

- *la suite de variables aléatoires (X_n)* ***converge en probabilité*** *vers X si, pour tout $\varepsilon > 0$,*

$$\mathbb{P}\left(|X_n - X| > \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0;$$

- *la suite de variables aléatoires (X_n)* ***converge presque sûrement*** *vers X si l'événement $\{X_n$ converge vers $X\}$ est presque sûr, cad si*

$$\mathbb{P}\left(\{\omega \in \Omega : \forall \varepsilon > 0, \exists N > 0, \forall n \geq N, |X_n(\omega) - X(\omega)| \leq \varepsilon\}\right) = 1.$$

Proposition 5.1.2. *Si une suite de variables aléatoires réelles (X_n) converge presque sûrement vers X, alors elle converge en probabilité vers X.*

La réciproque de la proposition ci-dessus n'est pas vraie comme en témoigne le contre-exemple suivant.

Contre-exemple 5.1.3. Soit (X_n) une suite de variables aléatoires indépendantes telles que $\mathbb{P}(X_n = 1) = \frac{1}{n}$ et $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$. Alors la suite (X_n) converge en probabilité vers 0 mais ne converge pas presque sûrement vers 0.

Convergence en loi

Définition 5.1.4. *Soient (X_n) et X des variables aléatoires réelles. On dit que la suite (X_n)* ***converge en loi*** *vers X si, pour tout $x \in \mathbf{R}$,*

$$\mathbb{P}(X_n \leq x) \underset{n\to\infty}{\longrightarrow} \mathbb{P}(X \leq x).$$

Proposition 5.1.5. *Soit (X_n) une suite de variables aléatoires à valeurs dans un ensemble dénombrable E, et soit X une variable aléatoire à valeurs dans E. Les assertions suivantes sont équivalentes :*

(i) X_n converge en loi vers X ;

(ii) pour tout $x \in E$, on a $\mathbb{P}(X_n = x) \underset{n\to\infty}{\longrightarrow} \mathbb{P}(X = x)$.

Proposition 5.1.6. *Si une suite de variables aléatoires réelles (X_n) converge en probabilité vers X, alors elle converge en loi vers X.*

La réciproque de la proposition ci-dessus n'est pas vraie comme en témoigne le contre-exemple suivant.

Contre-exemple 5.1.7. Prenons $\Omega = \{0,1\}$, $\mathbb{P} = \frac{1}{2}(\delta_0 + \delta_1)$ et $X_n(0) = 1$, $X_n(1) = 0$, $X(0) = 0$ et $X(1) = 1$. La suite de variables aléatoires (X_n) converge en loi vers X mais ne converge pas en probabilité vers X.

5.2 Loi faible des grands nombres

Théorème 5.2.1. *(Loi faible des grands nombres) Soit (X_n) une suite de variables aléatoires indépendantes, de même loi et admettant un moment d'ordre 2. Posons $m = \mathbb{E}[X_1]$ et $S_n = \sum_{k=1}^n X_k$. Alors $\frac{S_n}{n}$ converge en probabilité vers m, cad pour tout $\varepsilon > 0$,*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| > \varepsilon\right) \underset{n\to\infty}{\longrightarrow} 0.$$

Informellement, la loi faible des grands nombres signifie que la moyenne empirique converge vers la moyenne théorique, cad l'espérance. Le mot "faible" vient du fait qu'il existe une "loi forte des grands nombres" pour laquelle la convergence est plus précise (il s'agit d'une convergence presque sûre, et plus d'une convergence en probabilité) avec des hypothèses moins restrictives. La loi forte sera abordée en L3.

Exemple 5.2.2. Soit (X_n) une suite de variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $p \in [0,1]$. Posons $S_n = X_1 + \cdots + X_n$. Alors $\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} p$ en probabilité.

D'après l'exemple ci-dessus, si on lance un grand nombre de fois une pièce qui a une probabilité p de tomber sur pile, alors la proportion (empirique) de fois que l'on tombe sur pile après ces lancers est proche de la probabilité p (théorique) de tomber sur pile.

5.3 Lemmes de Borel-Cantelli

Lorsque (A_n) est une suite d'événements, la notation $\limsup A_n$ désigne l'événement :

$$\limsup A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k.$$

En particulier, $\limsup A_n$ est l'ensemble des $\omega \in \Omega$ tel qu'il existe une infinité d'indices k tels que $\omega \in A_k$.

Théorème 5.3.1. *(Premier lemme de Borel-Cantelli) Soit (A_n) une suite d'événements telle que $\sum_{n\geq 1} \mathbb{P}(A_n) < \infty$. Alors $\mathbb{P}(\limsup A_n) = 0$.*

En d'autres termes, si la série de terme général $\mathbb{P}(A_n)$ converge, alors
- presque sûrement, A_n n'est réalisé que pour un nombre fini d'indices n ;
- presque sûrement, A_n n'est pas réalisé à partir d'un certain rang ;
- avec probabilité 0, A_n est réalisé pour une infinité d'indices n.

Nous savons que la convergence presque sûre entraine la convergence en probabilité. Le résultat suivant donne une sorte de réciproque.

Corollaire 5.3.2. Soit (X_n) une suite de variables aléatoires et X une variable aléatoire. Supposons que, pour tout $\varepsilon > 0$, on ait $\sum_{n\geq 1} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$. Alors (X_n) converge presque sûrement vers X.

Théorème 5.3.3. *(Second lemme de Borel-Cantelli) Soit (A_n) une suite d'événements indépendants telle que $\sum_{n\geq 1} \mathbb{P}(A_n) = \infty$. Alors $\mathbb{P}(\limsup A_n) = 1$.*

En d'autres termes, si les événements (A_n) sont indépendants et si la série de terme général $\mathbb{P}(A_n)$ diverge, alors
- presque sûrement, A_n est réalisé pour une infinité d'indices n ;
- avec probabilité 0, A_n n'est pas réalisé à partir d'un certain rang.

5.4 Théorème central limite

Théorème 5.4.1. *(Théorème central limite) Soit (X_n) une suite de variables aléatoires indépendantes, de même loi et admettant un moment d'ordre 2. Posons $m = \mathbb{E}[X_1]$, $\sigma^2 = \mathbb{V}[X_1]$ et $S_n = \sum_{k=1}^{n} X_k$. Alors, pour tout $x \in \mathbf{R}$,*

$$\mathbb{P}\left(\sqrt{\frac{n}{\sigma^2}}\left(\frac{S_n}{n} - m\right) \leq x\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathrm{d}t.$$

Le théorème central limite doit son nom au fait qu'il joue un rôle fondamental en probabilités et en statistique, en particulier pour fournir des intervalles de confiance et des zones de rejets lors de tests statistiques. Il s'énoncera plus simplement en L3 en termes de convergence en loi. Plus précisément, la notion de convergence en loi telle que nous l'avons vue dans la définition 5.1.4 pour les variables aléatoires discrètes s'étend à des variables aléatoires réelles (indépendantes, avec un moment d'ordre 2) pas nécessairement discrètes. Le théorème central limite garantit que :

$$\sqrt{\frac{n}{\sigma^2}}\left(\frac{S_n}{n} - m\right) \overset{loi}{\to} U.$$

La quantité U désigne une variable aléatoire qui suit une loi fondamentale des probabilités : la *loi normale centrée réduite* (appelée également *loi gaussienne*). Cette variable aléatoire est telle que

$$\mathbb{P}(U \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathrm{d}t.$$

Informellement, il est *essentiel* de comprendre que :

- d'après la loi des grands nombres,

$$\frac{S_n}{n} \simeq m \quad \text{(en probabilité, presque sûrement);}$$

- d'après le théorème central limite,

$$\frac{S_n}{n} \simeq m + \sqrt{\frac{\sigma^2}{n}} U \quad \text{(en loi).}$$

L'essentiel

- savoir montrer qu'une suite de variables aléatoires convergent en probabilité ou convergent presque sûrement ;
- connaitre la loi faible des grands nombres ;
- appliquer les lemmes de Borel-Cantelli.