

VADE-MECUM DE PROBABILITÉS POUR L'AGRÉGATION INTERNE

Baptiste GORIN

4 août 2011

TABLE DES MATIÈRES

Table des matières	1
I Combinatoire et dénombrement	3
I.1 Cardinaux	3
I.2 Dénombrement	4
I.2.1 p -listes	4
I.2.2 Arrangements	4
I.2.3 Permutations	4
I.2.4 Combinaisons	5
I.2.5 Combinaisons avec répétition	6
I.3 Exemples fondamentaux	6
II Espace probabilisés	7
II.1 Espaces probabilisables	7
II.2 Espaces probabilisés	8
II.3 Probabilités conditionnelles	9
III Variables aléatoires réelles	11
III.1 Définitions	11
III.2 Variables aléatoires réelles discrètes	12
III.2.1 Définitions	12
III.2.2 Moments d'une variable aléatoire réelle discrète	12
III.2.3 Fonction génératrice d'une variable aléatoire réelle à valeurs dans \mathbb{N}	14
III.3 Lois usuelles discrètes	15
III.3.1 Loi uniforme	15
III.3.2 Loi de Bernoulli	16
III.3.3 Loi binomiale	16
III.3.4 Loi hypergéométrique	17
III.3.5 Loi géométrique	17
III.3.6 Loi de Poisson	18
III.4 Variables aléatoires réelles à densité	18
III.4.1 Définitions	18
III.4.2 Moments d'une variable aléatoire réelle à densité	19
III.5 Lois usuelles à densité	20
III.5.1 Loi uniforme	20
III.5.2 Loi exponentielle	20
III.5.3 Loi de Cauchy	21
III.5.4 Loi normale	21
III.5.5 Loi Gamma	22

IV Vecteurs aléatoires	24
IV.1 Généralités	24
IV.1.1 Définitions	24
IV.1.2 Indépendance de variables aléatoires	24
IV.1.3 Covariance	25
IV.2 Vecteurs aléatoires discrets	26
IV.3 Vecteurs aléatoires à densité	27
IV.4 Vecteurs gaussiens	29
V Théorèmes limites	30
V.1 Quelques résultats	30
V.1.1 Inégalités de Markov et de Bienaymé-Tchebychev	30
V.1.2 Lemme de Borel-Cantelli	30
V.2 Convergence des variables aléatoires	31
V.2.1 Différents modes de convergence	31
V.2.2 Comparaison des différents modes de convergence	31
V.3 Lois des grands nombres	32
V.3.1 Loi faible des grands nombres	32
V.3.2 Loi forte des grands nombres	32
V.4 Approximations	32
V.4.1 Approximation de la loi binomiale par la loi de Poisson	32
V.4.2 Approximation de la loi binomiale par la loi normale	33
V.5 Théorème limite central	33
Bibliographie	34

I.1 Cardinaux

Définition I.1.1. — Un ensemble E est dit fini s'il est vide ou s'il existe $n \in \mathbb{N}^*$ tel que E soit en bijection avec $\llbracket 1, n \rrbracket$.

L'entier n est le cardinal de E , noté $\text{Card}(E)$ ou $|E|$.

Proposition I.1.2. — Soient A un ensemble fini et B un ensemble tel qu'il existe une bijection entre A et B . Alors l'ensemble B est fini et $\text{Card}(B) = \text{Card}(A)$.

Définition I.1.3. — Deux ensembles A et B sont dits équipotents s'il existe une bijection entre A et B .

Proposition I.1.4. — Soient E un ensemble fini et $A \in \mathcal{P}(E)$. On a :

$$\text{Card}(A) \leq \text{Card}(E)$$

avec égalité si et seulement si $A = E$.

Théorème I.1.5. — Soit E un ensemble fini de cardinal n . On a :

$$\text{Card}(\mathcal{P}(E)) = 2^n.$$

Proposition I.1.6. — Soient E un ensemble fini et $A, B \in \mathcal{P}(E)$ deux sous-ensembles disjoints. On a :

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B).$$

Corollaire I.1.7. — Soient E un ensemble fini et $A_1, \dots, A_n \in \mathcal{P}(E)$ des sous-ensembles deux à deux disjoints. On a :

$$\text{Card}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \text{Card}(A_k).$$

En particulier, si A_1, \dots, A_n constituent une partition de E , on a :

$$\text{Card}(E) = \sum_{k=1}^n \text{Card}(A_k).$$

Corollaire I.1.8. — Soient E un ensemble fini et $A, B \in \mathcal{P}(E)$. On a :

$$\text{Card}(B \setminus A) = \text{Card}(B) - \text{Card}(A \cap B).$$

Corollaire I.1.9. — Soient E un ensemble fini et $A \in \mathcal{P}(E)$. On a :

$$\text{Card}(\overline{A}) = \text{Card}(E) - \text{Card}(A).$$

Théorème I.1.10. — Soient E un ensemble fini et $A, B \in \mathcal{P}(E)$. On a :

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B).$$

Théorème (formule du crible de Poincaré¹) I.1.11. — Soient E un ensemble fini et $A_1, \dots, A_n \in \mathcal{P}(E)$. On a :

$$\text{Card} \left(\bigcup_{k=1}^n A_k \right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \text{Card}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}).$$

Proposition I.1.12. — Soient E et F deux ensembles finis. On a :

$$\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F).$$

Corollaire I.1.13. — Soient E_1, \dots, E_n des ensembles finis. On a :

$$\text{Card}(E_1 \times \dots \times E_n) = \text{Card}(E_1) \times \dots \times \text{Card}(E_n).$$

I.2 Dénombrement

I.2.1 p -listes

Définition I.2.1. — Soient E un ensemble fini non vide et $p \in \mathbb{N}^*$. Une p -liste de E est un élément de E^p , c'est à dire un élément de la forme (a_1, \dots, a_p) avec $a_k \in E$ pour tout $k \in \llbracket 1; p \rrbracket$.

Théorème I.2.2. — Soient E un ensemble fini de cardinal n et $p \in \mathbb{N}^*$. Le nombre de p -listes de E est égal à n^p .

Définition I.2.3. — Soient E et F deux ensembles. L'ensemble des applications de E dans F se note F^E .

Théorème I.2.4. — Soient E et F deux ensembles finis. On a :

$$\text{Card}(F^E) = \text{Card}(F)^{\text{Card}(E)}$$

Remarque I.2.5. — Une p -liste de E peut être également vue comme une application de $\llbracket 1; p \rrbracket$ dans E .

I.2.2 Arrangements

Définition I.2.6. — Soient E un ensemble fini de cardinal n et $p \in \llbracket 1; n \rrbracket$. Un arrangement de p éléments de E est une p -liste de E constituée d'éléments deux à deux distincts, c'est-à-dire une application injective de $\llbracket 1; p \rrbracket$ dans E .

On note A_n^p le nombre d'arrangements de p éléments de E .

Remarque I.2.7. — Si on impose à une p -liste de E de ne contenir que des éléments deux à deux distincts, on ne peut espérer que la liste contienne plus de n éléments. Nécessairement, $p \leq n$.

Par convention, on pose alors $A_n^p = 0$ pour tout $p > n$.

Théorème I.2.8. — Soient E un ensemble de cardinal n et $p \in \llbracket 1; n \rrbracket$. Le nombre d'arrangements de p éléments de E est égal à

$$A_n^p = n \times (n - 1) \times \dots \times (n - p + 1).$$

I.2.3 Permutations

Définition I.2.9. — Soient E un ensemble de cardinal n . Une permutation de E est un arrangement de n éléments de E .

Définition I.2.10. — Soit $n \in \mathbb{N}$. On appelle « factorielle n » le nombre, noté $n!$, égal à

$$n! = 1 \times 2 \times 3 \times \dots \times n.$$

Par convention, on pose $0! = 1$.

Théorème I.2.11. — Soient E un ensemble de cardinal n . Le nombre de permutations de E est égal à $n!$.

Corollaire I.2.12. — Soient $n \in \mathbb{N}^*$ et $p \in \llbracket 1; n \rrbracket$. On a :

$$A_n^p = \frac{n!}{(n-p)!}.$$

1. Henri Poincaré, Nancy 1854 - Paris 1912

I.2.4 Combinaisons

Définition I.2.13. — Soient E un ensemble de cardinal n et $p \in \llbracket 0; n \rrbracket$. Une combinaison de p éléments de E est un sous-ensemble de E ayant p éléments.

On note $\binom{n}{p}$ le nombre de combinaisons de p éléments parmi n .

Remarque I.2.14. — La notation $\binom{n}{p}$ se lit « p parmi n ».

Les nombres $\binom{n}{p}$ sont également appelés coefficients binomiaux.

Remarque I.2.15. — Tout sous-ensemble de E comporte au plus n éléments. Nécessairement, $p \leq n$. Par convention, on pose alors $\binom{n}{p} = 0$ si $p > n$.

On convient de même que, si p ou n est négatif, $\binom{n}{p} = 0$.

Proposition I.2.16. — Soient $n, p \in \mathbb{N}$ avec $p \in \llbracket 0; n \rrbracket$. On a :

$$\binom{n}{0} = 1, \quad \binom{n}{1} = n, \quad \binom{n}{n} = 1, \quad \binom{n}{p} = \binom{n}{n-p}.$$

Proposition I.2.17. — Soient $n, p \in \mathbb{N}^*$ avec $p \in \llbracket 1; n \rrbracket$. On a :

$$p \binom{n}{p} = n \binom{n-1}{p-1}.$$

Théorème I.2.18. — Soient $n, p \in \mathbb{N}^*$ avec $p \in \llbracket 1; n \rrbracket$. On a :

$$\binom{n}{p} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}.$$

Théorème I.2.19. — Soient $n, p \in \mathbb{N}$ avec $p \in \llbracket 1; n \rrbracket$. On a :

$$\binom{n+1}{p} = \binom{n}{p} + \binom{n}{p-1} \quad (\text{formule de Pascal}^2)$$

Remarque (triangle de Pascal) I.2.20. — On range les coefficients binomiaux $\binom{n}{p}$ dans un tableau triangulaire tel que $\binom{n}{p}$ appartienne à la ligne d'indice n et la colonne d'indice p .

Les éléments de la $(n+1)$ -ième ligne sont calculés à partir de ceux de la n -ième ligne.

$n \backslash p$	0	1	2	3	4
0	1							
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
⋮	⋮					⋱		
⋮							$\binom{n}{p-1}$	$\binom{n}{p}$
⋮								$\binom{n+1}{p}$

Théorème (formule du binôme de Newton³) I.2.21. — Soient $a, b \in \mathbb{R}$ et $n \in \mathbb{N}$. On a :

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Proposition (formule de Vandermonde⁴) I.2.22. — Soient $m, n, p \in \mathbb{N}$ avec $p \leq m + n$. On a :

$$\sum_{k=0}^p \binom{m}{k} \binom{n}{p-k} = \binom{m+n}{p}.$$

2. Blaise Pascal, Clermont-Ferrand 1623 - Paris 1662
 3. Isaac Newton, Woolsthorpe 1642 - Kensington 1727
 4. Alexandre Vandermonde, Paris 1735 - Paris 1796

I.2.5 Combinaisons avec répétition

Définition I.2.23. — Soient E un ensemble de cardinal n et $p \in \mathbb{N}^*$. Une combinaison avec répétition de p éléments de E est une application $f : E \rightarrow \llbracket 0, p \rrbracket$ telle que

$$\sum_{x \in E} f(x) = p.$$

On note Γ_n^p le nombre de combinaisons avec répétition de p éléments pris parmi n .

Remarque I.2.24. — Si $E = \{e_1, \dots, e_n\}$. Une combinaison avec répétition de p éléments de E est une application $f : E \rightarrow \llbracket 0, p \rrbracket$ telle que

$$\sum_{k=1}^n f(e_k) = p.$$

où $f(e_k)$ est le nombre de fois que l'élément e_k est choisi.

Remarque I.2.25. — La notation Γ_n^p se lit « gamma n p ».

Théorème I.2.26. — Soient $n, p \in \mathbb{N}^*$. On a :

$$\Gamma_n^p = \binom{n+p-1}{p}$$

I.3 Exemples fondamentaux

Exemple I.3.1. — Soit N le nombre de tirages de p boules dans une urne en contenant n .

- Si les tirages sont ordonnés et successifs avec remise, alors on forme des p -listes ; d'où $N = n^p$.
- Si les tirages sont non ordonnés et successifs avec remise, alors on forme des combinaisons avec répétitions ; d'où $N = \Gamma_n^p$.
- Si les tirages sont ordonnés et successifs sans remise ($p \leq n$), alors on forme des arrangements ; d'où $N = A_n^p$.
- Si les tirages sont non ordonnés et successifs sans remise ou simultanés, alors $N = \binom{n}{p}$.

Type de tirage	Ordre	Répétitions d'éléments	Dénombrement
Successif avec remise	Ordonné	Oui	n^p p -listes
Successif avec remise	Non ordonné	Oui	Γ_n^p combinaisons avec répétitions
Successif sans remise	Ordonné	Non	A_n^p arrangements
Simultané	Non ordonné	Non	$\binom{n}{p}$ combinaisons

Exemple I.3.2. — Soit N le nombre de tirages de p boules dans une urne contenant n_1 boules de couleurs c_1, \dots, n_k boules de couleur c_k .

On cherche le nombre de tirages avec p_1 boules de couleur c_1, \dots, p_k boules de couleur c_k .

- Si les tirages sont non ordonnés et successifs sans remise ou simultanés, alors $N = \binom{n_1}{p_1} \binom{n_2}{p_2} \dots \binom{n_k}{p_k}$.
- Si les tirages sont ordonnés et successifs avec remise, alors $N = \binom{p}{p_1} \binom{p-p_1}{p_2} \dots \binom{p-p_1-\dots-p_{k-1}}{p_k} n_1^{p_1} n_2^{p_2} \dots n_k^{p_k}$.

Remarque I.3.3. — Lorsqu'un ensemble à dénombrer est défini à l'aide de la locution « au moins » il est généralement plus simple de dénombrer son complémentaire.

II.1 Espaces probabilisables

Définition II.1.1. — L'ensemble des résultats possibles d'une expérience aléatoire est appelé l'univers (ou l'univers des possibles ou l'ensemble fondamental) et noté Ω .

Tout élément de cet ensemble Ω est appelé un possible ou une éventualité. Les éléments de Ω sont notés ω .

Définition II.1.2. — Soit Ω l'univers correspondant à une expérience aléatoire. Un événement est une partie de Ω , c'est à dire un élément de $\mathcal{P}(\Omega)$.

Définition II.1.3. — Ω est un événement de Ω appelé événement certain de Ω .

\emptyset est un événement de Ω appelé événement impossible de Ω .

Un événement contenant un seul élément, si c'est possible, est appelé événement élémentaire de Ω .

Définition II.1.4. — Soit Ω un univers. Pour tout événement A de Ω , le complémentaire de A , noté \bar{A} , est appelé l'événement contraire de A .

Définition II.1.5. — Pour tous événements A et B d'un univers Ω tels que $A \subset B$, on dit que l'événement A implique l'événement B ; autrement dit, si A est réalisé, alors B est également réalisé.

Définition II.1.6. — Soient Ω un univers, A et B deux événements de Ω .

La réunion des événements A et B , notée $A \cup B$, est l'événement « A est réalisé ou B est réalisé ».

L'intersection des événements A et B , notée $A \cap B$, est l'événement « A est réalisé et B est réalisé ».

Remarque II.1.7. — Tous événements A et B de Ω vérifient les relations suivantes :

$$A \cap B \subset A \subset A \cup B \quad A \cap B \subset B \subset A \cup B \\ A \cup (A \cap B) = A \quad A \cap (A \cup B) = A \quad B \cup (A \cap B) = B \quad B \cap (A \cup B) = B$$

ainsi que les lois de Morgan :

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$$

Définition II.1.8. — Soient Ω un univers et A_1, A_2, \dots, A_n des événements .

• La réunion $\bigcup_{i=1}^n A_i$ est l'événement de Ω « l'un au moins des événements $A_k, 1 \leq k \leq n$, est réalisé ».

• L'intersection $\bigcap_{i=1}^n A_i$ est l'événement de Ω « tous les événements $A_k, 1 \leq k \leq n$, sont réalisés ».

Définition II.1.9. — Soient Ω un univers, A et B deux événements de Ω .

Si l'intersection de A et B est l'événement impossible :

$$A \cap B = \emptyset$$

alors les événements A et B sont dits incompatibles.

Définition II.1.10. — Soit Ω un univers. Tout sous-ensemble \mathcal{T} de $\mathcal{P}(\Omega)$ vérifiant :

- Ω appartient à \mathcal{T} ,
- \mathcal{T} est stable par passage au complémentaire; autrement dit, pour tout élément A de \mathcal{T} , \bar{A} appartient à \mathcal{T} ,
- \mathcal{T} est stable par réunion au plus dénombrable (σ -additivité); autrement dit, pour toute partie I de \mathbb{N} et toute famille $(A_i)_{i \in I}$ d'éléments de \mathcal{T} , $\bigcup_{i \in I} A_i$ appartient à \mathcal{T} ,

est appelé une tribu de Ω .

Tout élément de la tribu \mathcal{T} est appelé un événement de Ω .

Définition II.1.11. — Tout univers Ω muni d'une tribu \mathcal{T} est appelé espace probabilisable et noté (Ω, \mathcal{T}) .

Définition II.1.12. — Soient (Ω, \mathcal{T}) un espace probabilisable et $(A_i)_{i \in I}$ une famille d'événements de Ω , où I est une partie de \mathbb{N} . On dit que $(A_i)_{i \in I}$ est un système complet d'événements si

- $\bigcup_{i \in I} A_i = \Omega$;
- pour tous $i, j \in I, i \neq j$, on a $A_i \cap A_j = \emptyset$.

II.2 Espaces probabilisés

Définition II.2.1. — Une probabilité sur un espace probabilisable (Ω, \mathcal{T}) est une application $P : \mathcal{T} \rightarrow [0; 1]$ telle que :

- $P(\Omega) = 1$;
- pour toute famille $(A_i)_{i \in I}$ d'événements deux à deux incompatibles où I est une partie de \mathbb{N} , la série $\sum P(A_i)$ converge et on a

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i)$$

Pour tout événement A de \mathcal{T} , $P(A)$ est la probabilité de l'événement A .

Définition II.2.2. — Un espace probabilisable (Ω, \mathcal{T}) sur lequel est définie une probabilité P est appelé espace probabilisé et noté (Ω, \mathcal{T}, P) .

Proposition II.2.3. — Soit (Ω, \mathcal{T}, P) un espace probabilisé. La probabilité de l'événement impossible est nulle :

$$P(\emptyset) = 0.$$

Proposition II.2.4. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et A_1, A_2, \dots, A_n des événements deux à deux incompatibles. Alors on a :

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

Proposition II.2.5. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements. On a :

$$P(B \setminus A) = P(B) - P(A \cap B).$$

Proposition II.2.6. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et A un événement. On a :

$$P(\bar{A}) = 1 - P(A)$$

Proposition II.2.7. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements vérifiant $A \subset B$. Alors on a :

$$P(A) \leq P(B)$$

Proposition II.2.8. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements. On a :

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

Proposition II.2.9. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ un système complet d'événements de Ω où I est une partie de \mathbb{N} . Alors on a :

$$\sum_{i \in I} P(A_i) = 1.$$

Proposition II.2.10. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ un système complet d'événements de Ω où I est une partie de \mathbb{N} . Alors on a :

$$P(B) = \sum_{i \in I} P(B \cap A_i)$$

Théorème (inégalité de Boole¹) II.2.11. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ une famille d'événements de Ω où I est une partie de \mathbb{N} . Alors on a :

$$P\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} P(A_i)$$

1. George Boole, Lincoln 1815 - Balling Temple 1864

Théorème (formule du crible de Poincaré) II.2.12. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et A_1, A_2, \dots, A_n des événements. On a :

$$P\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}).$$

Théorème (propriété de la limite monotone) II.2.13. — Soit (Ω, \mathcal{T}, P) un espace probabilisé. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ croissante d'éléments de \mathcal{T} , c'est à dire vérifiant $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$, on a :

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

Pour toute suite $(A_n)_{n \in \mathbb{N}}$ décroissante d'éléments de \mathcal{T} , c'est à dire vérifiant $A_{n+1} \subset A_n$ pour tout $n \in \mathbb{N}$, on a :

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

Corollaire II.2.14. — Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'un espace probabilisé (Ω, \mathcal{T}, P) , on a :

$$\begin{aligned} P\left(\bigcup_{n=0}^{+\infty} A_n\right) &= \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=0}^n A_k\right) \\ P\left(\bigcap_{n=0}^{+\infty} A_n\right) &= \lim_{n \rightarrow +\infty} P\left(\bigcap_{k=0}^n A_k\right). \end{aligned}$$

Théorème II.2.15. — Soit $(\Omega, \mathcal{P}(\Omega))$ un espace probabilisable fini. Il existe une probabilité P , unique, prenant la même valeur sur tous les événements élémentaires :

$$\forall \omega \in \Omega, \quad P(\{\omega\}) = \frac{1}{\text{Card}(\Omega)}.$$

Alors, pour tout événement A de $\mathcal{P}(\Omega)$, on a : $P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$.

La probabilité ainsi définie est appelée probabilité uniforme sur Ω . On dit qu'il y a équiprobabilité.

Remarque II.2.16. — La formule $P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$ s'énonce également : la probabilité de l'événement A est le quotient du nombre des cas favorables (c'est à dire réalisant l'événement A) sur le nombre de cas possibles.

Remarque II.2.17. — Si l'univers Ω est muni de la probabilité uniforme, les calculs des probabilités se ramènent à des calculs de dénombrement.

II.3 Probabilités conditionnelles

Théorème - Définition II.3.1. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et A un événement tel que $P(A) \neq 0$. L'application

$$P_A : \begin{cases} \mathcal{T} & \longrightarrow \mathbb{R} \\ B & \longmapsto \frac{P(B \cap A)}{P(A)} \end{cases}$$

est une probabilité sur (Ω, \mathcal{T}, P) , appelée la probabilité conditionnelle sachant A (ou la probabilité conditionnelle relative à A).

Pour tout événement B , $P_A(B)$ est appelé la probabilité de B sachant A .

Remarque II.3.2. — Au lieu de $P_A(B)$, on note aussi $P(B|A)$.

Théorème (formule des probabilités composées) II.3.3. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et (A_1, A_2, \dots, A_n) une famille d'événements telle que $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0$. Alors on a :

$$P(A_1 \cap A_2 \cap \dots \cap A_{n-1} \cap A_n) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap A_2 \cap \dots \cap A_{n-1}}(A_n).$$

Remarque II.3.4. — Cette formule est souvent employée lorsque les événements A_1, A_2, \dots, A_n sont dans un ordre chronologique.

Théorème (formule des probabilités totales) II.3.5. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ un système complet d'événements (où I est une partie de \mathbb{N}) tels que, pour tout $i \in I$, $P(A_i) \neq 0$. Pour tout événement B , on a :

$$P(B) = \sum_{i \in I} P(A_i)P_{A_i}(B)$$

Remarque II.3.6. — Cette formule est souvent employée quand une expérience se déroule en plusieurs étapes et que la première aboutit à plusieurs résultats incompatibles entre eux et donnant un système complet d'événements.

Théorème (formule de Bayes² ou de probabilité des causes) II.3.7. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements tels que $P(A)P(B) \neq 0$. Alors on a :

$$P_B(A) = \frac{P(A)P_A(B)}{P(B)}.$$

Remarque II.3.8. — Cette formule permet, d'une certaine façon, de « remonter le temps » : en effet, si l'événement B se produit à une date postérieure à celle de A , elle permet de déduire, de $P_A(B)$ qui respecte la chronologie, la probabilité $P_B(A)$ qui, elle, remonte cette chronologie.

Corollaire II.3.9. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, (A_1, A_2, \dots, A_n) un système complet d'événements et B un événement tels que $P(A_k) \neq 0$ pour tout $k \in \llbracket 1; n \rrbracket$ et $P(B) \neq 0$. Alors on a :

$$P_B(A_k) = \frac{P_{A_k}(B)P(A_k)}{\sum_{i=1}^n P_{A_i}(B)P(A_i)}.$$

Définition II.3.10. — Soient (Ω, \mathcal{T}, P) un espace probabilisé. Deux événements A et B sont dits indépendants si

$$P(A \cap B) = P(A)P(B)$$

Théorème II.3.11. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements avec $P(A) \neq 0$. Les événements A et B sont indépendants si et seulement si on a

$$P_A(B) = P(B)$$

Proposition II.3.12. — Soient (Ω, \mathcal{T}, P) un espace probabilisé, A et B deux événements indépendants. Alors les événements A et \overline{B} , les événements \overline{A} et B et les événements \overline{A} et \overline{B} sont indépendants.

Définition II.3.13. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ une famille d'événements de Ω où I est une partie de \mathbb{N} . Les événements $(A_i)_{i \in I}$ sont dits deux à deux indépendants si, pour tous $i, j \in I$, $i \neq j$, on a :

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

Définition II.3.14. — Soient (Ω, \mathcal{T}, P) un espace probabilisé et $(A_i)_{i \in I}$ une famille d'événements de Ω où I est une partie de \mathbb{N} . Les événements $(A_i)_{i \in I}$ sont dits mutuellement indépendants (ou indépendants dans leur ensemble) si, pour toute partie finie non vide J de I , on a :

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

Remarque II.3.15. — Si des événements sont mutuellement indépendants, alors ils sont deux à deux indépendants. La réciproque est fautive.

Proposition II.3.16. — Soit (Ω, \mathcal{T}, P) un espace probabilisé. Si des événements A_1, A_2, \dots, A_n sont mutuellement indépendants, alors les événements B_1, B_2, \dots, B_n , où B_k est A_k ou $\overline{A_k}$, sont mutuellement indépendants.

2. Thomas Bayes, Londres 1702 - Tunbridge Wells (Kent) 1761

III.1 Définitions

Définition III.1.1. — Soit (Ω, \mathcal{T}) un espace probabilisable. On appelle variable aléatoire réelle toute application mesurable de (Ω, \mathcal{T}) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, autrement dit toute application X de Ω dans \mathbb{R} telle que, pour tout intervalle I de \mathbb{R} , on ait :

$$X^{-1}(I) = \{\omega \in \Omega; X(\omega) \in I\} \in \mathcal{T}.$$

Proposition - Définition III.1.2. — Soit X une variable aléatoire définie sur un espace probabilisé (Ω, \mathcal{T}, P) à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

L'application $P_X : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ définie par :

$$P_X(B) = P(X^{-1}(B))$$

est une probabilité sur l'espace probabilisable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Cette probabilité est appelée loi de probabilité de la variable aléatoire X . On dit aussi que X suit la loi de probabilité P_X .

Définition III.1.3. — Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{T}, P) . La fonction de répartition de X est l'application

$$F_X : \begin{cases} \mathbb{R} & \rightarrow & \mathbb{R} \\ x & \mapsto & P_X(X \leq x) \end{cases}$$

Remarque III.1.4. — Pour tous $a, b \in \mathbb{R}$, on a : $P(a < X \leq b) = F_X(b) - F_X(a)$.

Proposition III.1.5. — La fonction de répartition F_X d'une variable aléatoire réelle X vérifie les propriétés suivantes :

- pour tout $x \in \mathbb{R}$, $F_X(x) \in [0, 1]$;
- on a :

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} F_X(x) = 1$$

- F_X est croissante sur \mathbb{R} , continue à droite et admet une limite à gauche en tout point de \mathbb{R} ; de plus, pour tout $x \in \mathbb{R}$, on a :

$$\lim_{t \rightarrow x, t < x} F_X(t) = F_X(x) - P(X = x) = P(X < x).$$

Corollaire III.1.6. — Pour toute variable aléatoire réelle X , d'un espace probabilisé (Ω, \mathcal{T}, P) , la fonction de répartition de X est continue en x si et seulement si $P(X = x) = 0$.

Proposition III.1.7. — Soient X et Y deux variables aléatoires réelles définies sur un espace probabilisé (Ω, \mathcal{T}, P) .

Alors $F_X = F_Y$ si et seulement si $P_X = P_Y$.

Ainsi la fonction de répartition caractérise la loi de probabilité

III.2 Variables aléatoires réelles discrètes

III.2.1 Définitions

Définition III.2.1. — Une variable aléatoire réelle X définie sur un espace probabilisable (Ω, \mathcal{T}) est dite discrète si son image $X(\Omega)$ est au plus dénombrable.

Proposition - Définition III.2.2. — Soit (Ω, \mathcal{T}) un espace probabilisable. Pour tout $a \in \mathbb{R}$, l'application :

$$X : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto a \end{cases}$$

est une variable aléatoire discrète finie appelée variable aléatoire constante ou certaine.

Proposition - Définition III.2.3. — Soit (Ω, \mathcal{T}) un espace probabilisable. Pour tout événement A de Ω , l'application :

$$X : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \end{cases}$$

est une variable aléatoire discrète finie appelée variable aléatoire indicatrice de l'événement A , notée χ_A .

Proposition - Définition III.2.4. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisable (Ω, \mathcal{T}) . Alors la famille $(X = x)_{x \in X(\Omega)}$ est un système complet d'événements de (Ω, \mathcal{T}) appelé le système complet d'événements associé à la variable aléatoire X .

Définition III.2.5. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, \mathcal{T}, P) . La loi de probabilité (ou distribution) de X est l'application

$$P_X : \begin{cases} X(\Omega) & \longrightarrow \mathbb{R} \\ x & \longmapsto P(X = x) \end{cases}$$

Autrement dit, la loi de probabilité de X est la donnée de $X(\Omega)$ et des probabilités $P(X = x)$ pour tout $x \in X(\Omega)$.

Proposition III.2.6. — Soient X une variable aléatoire réelle discrète définie sur un espace probabilisable (Ω, \mathcal{T}) et φ une application définie sur $X(\Omega)$. Alors l'application

$$Y : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto \varphi(X(\omega)) \end{cases}$$

est une variable aléatoire réelle discrète notée $\varphi(X)$.

La loi de Y est donnée par :

$$P_Y : \begin{cases} Y(\Omega) & \longrightarrow \mathbb{R} \\ y & \longmapsto \sum_{\substack{x \in X(\Omega) \\ \varphi(x)=y}} P(X = x) \end{cases}$$

Proposition III.2.7. — Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{T}, P) .

Alors la variable aléatoire X est discrète finie (resp. infinie) si et seulement si sa fonction de répartition F_X est une fonction en escalier dont l'ensemble des points de discontinuité est fini (resp. dénombrable). Cet ensemble est alors égal à $X(\Omega)$.

Remarque III.2.8. — Si $X(\Omega) \subset \mathbb{Z}$, on a, pour tout $k \in \mathbb{Z}$, $P(X = k) = F_X(k) - F_X(k - 1)$.

III.2.2 Moments d'une variable aléatoire réelle discrète

Définition III.2.9. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, \mathcal{T}, P) . Si la série de terme général $xP(X = x)$, avec $x \in X(\Omega)$, est absolument convergente, on dit que X admet une espérance $E(X)$ et on pose :

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

Remarque III.2.10. — Si $X(\Omega) = \mathbb{N}$, on a $E(X) = \sum_{n=0}^{+\infty} nP(X = n)$.

Proposition III.2.11. — Soit (Ω, \mathcal{T}, P) un espace probabilisé.

- Pour tout événement A , la variable aléatoire indicatrice de l'événement A admet une espérance et on a :

$$E(\chi_A) = P(A).$$

- Pour tout $a \in \mathbb{R}$, la variable aléatoire certaine égale à a admet une espérance et on a :

$$E(a) = a.$$

Proposition III.2.12. — Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{T}, P) , admettant une espérance.

- Si $X \geq 0$, alors $E(X) \geq 0$.
- Si, de plus, $E(X) = 0$, alors la variable aléatoire X est presque sûrement nulle.

Proposition III.2.13. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, \mathcal{T}, P) , admettant une espérance, et $\alpha, \beta \in \mathbb{R}$.

Alors la variable aléatoire $\alpha X + \beta$ admet une espérance et

$$E(\alpha X + \beta) = \alpha E(X) + \beta.$$

Proposition (linéarité de l'espérance) III.2.14. — Soient X et Y deux variables aléatoires réelles discrètes définies sur un espace probabilisé (Ω, \mathcal{T}, P) , admettant une espérance, et $\alpha, \beta \in \mathbb{R}$.

Alors la variable aléatoire $\alpha X + \beta Y$ admet une espérance et

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

Autrement dit, l'application $X \mapsto E(X)$ est une forme linéaire de l'espace vectoriel des variables aléatoires admettant une espérance.

Définition III.2.15. — Une variable aléatoire réelle admettant une espérance nulle est dite centrée.

Proposition - Définition III.2.16. — Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{T}, P) , admettant une espérance. $X - E(X)$ est une variable aléatoire réelle centrée appelée la variable aléatoire centrée associée à X .

Proposition III.2.17. — Soient X et Y deux variables aléatoires réelles discrètes définies sur un espace probabilisé (Ω, \mathcal{T}, P) , admettant une espérance.

Si $X \leq Y$, alors $E(X) \leq E(Y)$.

Théorème (de transfert) III.2.18. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, \mathcal{T}, P) et φ une application définie sur $X(\Omega)$.

La variable aléatoire réelle discrète $\varphi(X)$ admet une espérance si et seulement si la série $\sum_{x \in X(\Omega)} \varphi(x)P(X = x)$

est absolument convergente.

En cas de convergence absolue, on a alors :

$$E(\varphi(X)) = \sum_{x \in X(\Omega)} \varphi(x)P(X = x).$$

Définition III.2.19. — Soit $p \in \mathbb{R}$, $p \geq 1$. On note $L_d^p(\Omega, \mathcal{T}, P)$ (ou plus simplement L_d^p) l'ensemble des variables aléatoires réelles discrètes X définie sur l'espace probabilisé (Ω, \mathcal{T}, P) et telles que X^p admet une espérance. On dit alors que X admet un moment d'ordre p , noté $m_p(X)$, où :

$$m_p(X) = E(X^p) = \sum_{x \in X(\Omega)} x^p P(X = x).$$

Proposition III.2.20. — • Pour $p \geq 1$, L_d^p est un espace vectoriel.

- Soient $p, q \geq 1$ tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que p et q sont conjugués), $X \in L_d^p$ et $Y \in L_d^q$. Alors $XY \in L_d^1$ et :

$$E(|XY|) \leq (E(|X|^p))^{\frac{1}{p}} (E(|Y|^q))^{\frac{1}{q}}$$

- Soient $p, p' \geq 1$, $p' \leq p$. Si $X \in L_d^p$, alors $X \in L_d^{p'}$ et :

$$(E(|X|^{p'}))^{\frac{1}{p'}} \leq (E(|X|^p))^{\frac{1}{p}}$$

Proposition - Définition III.2.21. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2.

Alors la variable aléatoire $X - E(X)$ admet un moment d'ordre 2.

La variance de X , notée $\text{Var}(X)$, est le réel

$$\text{Var}(X) = m_2(X - E(X)) = E((X - E(X))^2).$$

Proposition III.2.22. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2.

Alors on a $\text{Var}(X) \geq 0$.

De plus, $\text{Var}(X) = 0$ si et seulement si $X = E(X)$ presque sûrement.

Proposition III.2.23. — Soient X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2, et $\alpha, \beta \in \mathbb{R}$.

Alors la variable aléatoire $\alpha X + \beta$ admet une variance et

$$\text{Var}(\alpha X + \beta) = \alpha^2 \text{Var}(X).$$

Proposition (formule de Kœnig¹-Huygens²) III.2.24. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2. On a :

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

Définition III.2.25. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2. L'écart-type de X , noté $\sigma(X)$, est le réel $\sigma(X) = \sqrt{\text{Var}(X)}$.

Définition III.2.26. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2. Si $E(X) = 0$ et $\sigma(X) = 1$, la variable aléatoire X est dite centrée réduite.

Proposition- Définition III.2.27. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant une variance non nulle. $X^* = \frac{X - E(X)}{\sigma(X)}$ est une variable aléatoire réelle discrète centrée réduite, appelée la variable aléatoire réelle centrée réduite associée à X .

III.2.3 Fonction génératrice d'une variable aléatoire réelle à valeurs dans \mathbb{N}

Définition III.2.28. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et telle que $X(\Omega) \subset \mathbb{N}$. La fonction génératrice de X est la série entière G_X définie par :

$$G_X(s) = E(s^X) = \sum_{n=0}^{+\infty} P(X = n)s^n.$$

Proposition III.2.29. — Le rayon de convergence de la série entière définissant G_X est supérieur ou égal à 1. On a :

$$\forall s \in [-1, 1], \quad |G_X(s)| \leq 1 \quad \text{et} \quad G_X(1) = 1$$

De plus, la fonction G_X est continue sur $[-1, 1]$ et de classe \mathcal{C}^∞ sur $] -1, 1[$.

Théorème (d'unicité) III.2.30. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et telle que $X(\Omega) \subset \mathbb{N}$.

La fonction génératrice de X caractérise la loi de X .

Pour tout $n \in \mathbb{N}$, on a :

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

Théorème III.2.31. — Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et telle que $X(\Omega) \subset \mathbb{N}$.

X admet un moment d'ordre $p \in \mathbb{N}^*$ si et seulement si sa fonction génératrice G_X admet une dérivée à gauche $G_X^{(p)}(1^-)$ d'ordre p .

On a alors :

$$G_X^{(p)}(1^-) = E(X(X-1) \cdots (X-p+1))$$

1. Samuel Kœnig, Büdingen 1712 - Zuilenstein 1757
 2. Christiaan Huygens, La Haye 1629 - La Haye 1695

En particulier, on a :

$$E(X) = G'_X(1^-) \quad \text{et} \quad \text{Var}(X) = G''_X(1^-) + G'_X(1^-) - (G'_X(1^-))^2.$$

Théorème III.2.32. — Soient X et Y deux variables aléatoires réelles discrètes indépendantes définies sur un espace probabilisé (Ω, \mathcal{T}, P) , à valeurs dans \mathbb{N} , de fonctions génératrices G_X et G_Y . Alors la fonction génératrice de $X + Y$ est donnée par :

$$G_{X+Y} = G_X G_Y$$

Corollaire III.2.33. — Soient X_1, \dots, X_n des variables aléatoires réelles discrètes définies sur un espace probabilisé (Ω, \mathcal{T}, P) , mutuellement indépendantes et à valeurs dans \mathbb{N} , de fonctions génératrices G_{X_1}, \dots, G_{X_n} . Alors la fonction génératrice de $S_n = \sum_{k=1}^n X_k$ est donnée par :

$$G_{S_n} = G_{X_1} G_{X_2} \cdots G_{X_n}$$

Corollaire III.2.34. — Soient X_1, \dots, X_n des variables aléatoires réelles discrètes définies sur un espace probabilisé (Ω, \mathcal{T}, P) , mutuellement indépendantes, identiquement distribuées et à valeurs dans \mathbb{N} , de fonction génératrice G .

Alors la fonction génératrice de $S_n = \sum_{k=1}^n X_k$ est donnée par :

$$G_{S_n} = G^n$$

III.3 Lois usuelles discrètes

III.3.1 Loi uniforme

Exemple (situation type) III.3.1. — Une urne contient n boules numérotées de 1 à n . Considérons l'expérience aléatoire qui consiste à tirer une boule et à noter son numéro X .

X est une variable aléatoire telle que $X(\Omega) = \llbracket 1; n \rrbracket$.

On se place dans le cadre de l'équiprobabilité. Ainsi, pour tout $k \in \llbracket 1; n \rrbracket$, on a :

$$P(X = k) = \frac{1}{n}$$

Définition III.3.2. — Une variable aléatoire réelle discrète X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi uniforme sur $\llbracket 1; n \rrbracket$, où $n \in \mathbb{N}^*$, si

$$\begin{cases} X(\Omega) = \llbracket 1; n \rrbracket \\ \forall k \in \llbracket 1; n \rrbracket, \quad P(X = k) = \frac{1}{n} \end{cases}$$

« X suit la loi uniforme sur $\llbracket 1; n \rrbracket$ » se note $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$.

Proposition III.3.3. — Toute variable aléatoire X de loi uniforme sur $\llbracket 1; n \rrbracket$ admet une espérance et une variance égales à :

$$E(X) = \frac{n+1}{2} \quad \text{et} \quad \text{Var}(X) = \frac{n^2-1}{12}$$

Définition III.3.4. — Une variable aléatoire réelle discrète X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi uniforme sur $\llbracket a; b \rrbracket$, où $a, b \in \mathbb{Z}$, $a < b$, si

$$\begin{cases} X(\Omega) = \llbracket a; b \rrbracket \\ \forall k \in \llbracket a; b \rrbracket, \quad P(X = k) = \frac{1}{b-a+1} \end{cases}$$

« X suit la loi uniforme sur $\llbracket a; b \rrbracket$ » se note $X \hookrightarrow \mathcal{U}(\llbracket a; b \rrbracket)$.

Proposition III.3.5. — Pour toute variable aléatoire X de loi uniforme sur $\llbracket a; b \rrbracket$ avec $a < b$, la variable aléatoire $Y = X - a + 1$ suit la loi uniforme sur $\llbracket 1; b - a + 1 \rrbracket$.

Proposition III.3.6. — Toute variable aléatoire X de loi uniforme sur $\llbracket a; b \rrbracket$ admet une espérance et une variance égales à :

$$E(X) = \frac{a+b}{2} \quad \text{et} \quad \text{Var}(X) = \frac{(b-a)(b-a+1)}{12}$$

III.3.2 Loi de Bernoulli

Schéma théorique

Considérons une expérience aléatoire ayant deux issues, à savoir : succès avec une probabilité p , échec avec une probabilité $q = 1 - p$.

Soit X la variable aléatoire valant 1 en cas de succès, 0 en cas d'échec.

On a alors $X(\Omega) = \{0; 1\}$ et

$$P(X = 1) = p, \quad P(X = 0) = q.$$

Définition III.3.7. — Soit $p \in]0; 1[$. Une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi de Bernoulli³ de paramètre p si $X(\Omega) = \{0; 1\}$ et $P(X = 1) = p$.

« X suit la loi de Bernoulli de paramètre p » se note $X \hookrightarrow \mathcal{B}(p)$.

Exemple (situation type) III.3.8. — On considère une urne contenant un nombre fini de boules blanches et de boules noires supposées indiscernables au toucher, la proportion de boules blanches étant p .

Considérons l'expérience aléatoire consistant à tirer une boule de l'urne.

On se place dans le cadre de l'équiprobabilité.

Soit X la variable aléatoire valant 1 si la boule tirée est blanche, 0 si elle est noire.

Alors X suit la loi de Bernoulli $\mathcal{B}(p)$.

Proposition III.3.9. — Soit X une variable aléatoire de Bernoulli de paramètre p . Alors X a une espérance et une variance égales à :

$$E(X) = p \quad \text{et} \quad \text{Var}(X) = p(1 - p) = pq$$

Proposition III.3.10. — Soit X une variable aléatoire suivant la loi de Bernoulli $\mathcal{B}(p)$. Sa fonction génératrice G_X est définie par :

$$G_X(s) = 1 - p + ps$$

Son rayon de convergence est $R = +\infty$.

III.3.3 Loi binomiale

Schéma théorique

Considérons une expérience aléatoire se déroulant en n épreuves mutuellement indépendantes. Chaque épreuve a deux issues, à savoir : succès avec une probabilité p , échec avec une probabilité $q = 1 - p$.

Soit X la variable aléatoire égale au nombre de succès à l'issue de cette série de n épreuves.

On a alors $X(\Omega) = \llbracket 0; n \rrbracket$ et

$$\forall k \in \llbracket 0; n \rrbracket, \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} = \binom{n}{k} p^k q^{n-k}.$$

Définition III.3.11. — Soient $n \in \mathbb{N}^*$ et $p \in]0; 1[$. Une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi binomiale de paramètre (n, p) si

$$\left\{ \begin{array}{l} X(\Omega) = \llbracket 0; n \rrbracket \\ \forall k \in \llbracket 0; n \rrbracket, \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} = \binom{n}{k} p^k q^{n-k} \end{array} \right.$$

« X suit la loi binomiale de paramètre (n, p) » se note $X \hookrightarrow \mathcal{B}(n, p)$.

Remarque III.3.12. — Une variable aléatoire qui suit la loi binomiale $\mathcal{B}(1, p)$ est une variable aléatoire de Bernoulli de paramètre p .

Exemple (situation type) III.3.13. — On considère une urne contenant un nombre fini de boules blanches et de boules noires supposées indiscernables au toucher, la proportion de boules blanches étant p .

Considérons l'expérience consistant à effectuer n tirages avec remise après chaque tirage.

On note X la variable aléatoire égale au nombre de boules blanches tirées.

Alors X suit la loi binomiale $\mathcal{B}(n, p)$.

Proposition III.3.14. — Soit X une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, p)$. Alors X a une espérance et une variance égales à :

$$E(X) = np \quad \text{et} \quad \text{Var}(X) = np(1 - p) = npq$$

3. Jakob Bernoulli, Bâle 1655 - Bâle 1705

Proposition III.3.15. — Soit X une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, p)$. Sa fonction génératrice G_X est définie par :

$$G_X(s) = (1 - p + ps)^n$$

Son rayon de convergence est $R = +\infty$.

Proposition III.3.16. — Soient X et Y deux variables aléatoires réelles indépendantes définies sur un espace probabilisé (Ω, \mathcal{T}, P) suivant les lois binomiales $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$ respectivement. Alors la variable aléatoire $X + Y$ suit la loi binomiale $\mathcal{B}(m + n, p)$.

Corollaire III.3.17. — Soient X_1, X_2, \dots, X_n n variables aléatoires réelles mutuellement indépendantes suivant la loi de Bernoulli $\mathcal{B}(p)$.

Alors la variable aléatoire $X_1 + X_2 + \dots + X_n$ suit la loi binomiale $\mathcal{B}(n, p)$.

III.3.4 Loi hypergéométrique

Schéma théorique

Soit E un ensemble de cardinal N dont a éléments de type 1 et $b = N - a$ éléments de type 2.

Considérons l'expérience consistant à tirer, sans remise, n éléments de E ($n \leq N$).

Soit X la variable aléatoire égale au nombre d'éléments de type 1 obtenus.

On a alors $X(\Omega) = [\max(0, n - b); \min(a, n)]$ et

$$P(X = k) = \frac{\binom{a}{k} \binom{b}{n-k}}{\binom{a+b}{n}} = \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}}$$

où $p = \frac{a}{N} = \frac{a}{a+b}$ est la proportion d'éléments de type 1 dans E .

Définition III.3.18. — Soient $n, N \in \mathbb{N}^*$, $n \leq N$, et $p \in]0, 1[$. Une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi hypergéométrique de paramètre (N, n, p) si

$$\left\{ \begin{array}{l} X(\Omega) = [\max(0, n - N(1-p)); \min(Np, n)] \\ \forall k \in X(\Omega), \quad P(X = k) = \frac{\binom{Np}{k} \binom{N(1-p)}{n-k}}{\binom{N}{n}} \end{array} \right.$$

« X suit la loi hypergéométrique de paramètre (N, n, p) » se note $X \hookrightarrow \mathcal{H}(N, n, p)$.

Proposition III.3.19. — Soit X une variable aléatoire suivant la loi hypergéométrique $\mathcal{H}(N, n, p)$. Alors X a une espérance et une variance égales à

$$E(X) = np \quad \text{et} \quad \text{Var}(X) = np(1-p) \frac{N-n}{N-1}$$

III.3.5 Loi géométrique

Schéma théorique

Considérons une expérience aléatoire se déroulant en une infinité d'épreuves mutuellement indépendantes.

Chaque épreuve a deux issues, à savoir : succès avec une probabilité p , échec avec une probabilité $q = 1 - p$.

Soit X la variable aléatoire égale au nombre d'épreuves nécessaires pour obtenir le premier succès.

On a alors $X(\Omega) = \mathbb{N}^*$ et

$$P(X = k) = p(1-p)^{k-1} = pq^{k-1} \text{ pour tout } k \in \mathbb{N}^*.$$

Définition III.3.20. — Soit $p \in]0, 1[$. Une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi géométrique de paramètre p si

$$\left\{ \begin{array}{l} X(\Omega) = \mathbb{N}^* \\ \forall k \in \mathbb{N}^*, \quad P(X = k) = p(1-p)^{k-1} = pq^{k-1} \end{array} \right.$$

« X suit la loi géométrique de paramètre p » se note $X \hookrightarrow \mathcal{G}(p)$.

Exemple (situation type) III.3.21. — On considère une urne contenant un nombre fini de boules blanches et de boules noires supposées indiscernables au toucher, la proportion de boules blanches étant p .

Considérons l'expérience consistant à effectuer une infinité de tirages avec remise après chaque tirage. On note X la variable aléatoire égale au nombre de tirages nécessaires pour obtenir la première boule blanche. Alors X suit la loi géométrique $\mathcal{G}(p)$.

Proposition III.3.22. — Soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$. Alors X a une espérance et une variance égales à :

$$E(X) = \frac{1}{p} \quad \text{et} \quad \text{Var}(X) = \frac{1-p}{p^2} = \frac{q}{p^2}.$$

Proposition III.3.23. — Soit X une variable aléatoire suivant la loi géométrique $\mathcal{G}(p)$. Sa fonction génératrice G_X est définie par :

$$G_X(s) = \frac{ps}{1-qs}$$

Son rayon de convergence est $R = \frac{1}{q}$.

III.3.6 Loi de Poisson

Définition III.3.24. — Soit $\lambda \in \mathbb{R}_+^*$. Une variable aléatoire X définie sur un espace probabilisé (Ω, \mathcal{T}, P) suit la loi de Poisson⁴ de paramètre λ si

$$\begin{cases} X(\Omega) = \mathbb{N} \\ \forall k \in \mathbb{N}, \quad P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \end{cases}$$

« X suit la loi de Poisson de paramètre λ » se note $X \hookrightarrow \mathcal{P}(\lambda)$.

Proposition III.3.25. — Soit X une variable aléatoire suivant la loi de Poisson $\mathcal{P}(\lambda)$. Alors X a une espérance et une variance égales à

$$E(X) = \lambda \quad \text{et} \quad \text{Var}(X) = \lambda.$$

Proposition III.3.26. — Soit X une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\lambda)$. Sa fonction génératrice G_X est définie par :

$$G_X(s) = e^{\lambda(s-1)}$$

Son rayon de convergence est $R = +\infty$.

Proposition III.3.27. — Soient X et Y deux variables aléatoires réelles indépendantes définies sur un même espace probabilisé (Ω, \mathcal{T}, P) suivant les lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$ respectivement. Alors la variable aléatoire $X + Y$ suit la loi de Poisson $\mathcal{P}(\lambda + \mu)$.

III.4 Variables aléatoires réelles à densité

III.4.1 Définitions

Définition III.4.1. — On appelle densité de probabilité toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}_+$ telle que $\int_{\mathbb{R}} f(t)dt = 1$.

Définition III.4.2. — Soit f une densité de probabilité. Une variable aléatoire réelle X possède la loi de densité f si, pour tout intervalle I de \mathbb{R} , on a :

$$P(X \in I) = \int_I f(t)dt$$

Définition III.4.3. — Soit X une variable aléatoire de densité f_X . La fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par $F(x) = P(X \leq x)$ est la fonction de répartition de X .

Proposition III.4.4. — Soit X une variable aléatoire de densité f_X et de fonction de répartition F_X . Alors F_X est continue en tout point de \mathbb{R} et on a $P(X = x) = 0$ pour tout $x \in \mathbb{R}$. De plus, pour tout $x \in \mathbb{R}$, on a :

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

et F_X est dérivable en tout point de continuité de f_X .

4. Siméon Denis Poisson, Pithiviers 1781 - Sceaux 1840

Remarque III.4.5. — Sous les hypothèses de la proposition, une densité f_X de X est donnée, par exemple, par :

$$f_X(x) = \begin{cases} F'_X(x) & \text{si } F_X \text{ est dérivable en } x \\ 0 & \text{sinon} \end{cases}$$

Théorème III.4.6. — Soient X une variable aléatoire réelle admettant une densité f_X et φ une fonction réelle définie sur \mathbb{R} strictement monotone et dérivable. La variable aléatoire réelle $Y = \varphi(X)$ admet une densité f_Y donnée par :

$$f_Y(y) = \begin{cases} f_X(\varphi^{-1}(y)) \left| (\varphi^{-1})'(y) \right| & \text{si } y \in \varphi(\mathbb{R}) \\ 0 & \text{sinon} \end{cases}$$

III.4.2 Moments d'une variable aléatoire réelle à densité

Définition III.4.7. — Soit X une variable aléatoire réelle de densité f_X . Si l'intégrale $\int_{-\infty}^{+\infty} t f_X(t) dt$ est absolument convergente, on dit que X admet une espérance $E(X)$ et on pose :

$$E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt.$$

Proposition III.4.8. — Soient X une variable aléatoire réelle à densité, admettant une espérance, et $\alpha, \beta \in \mathbb{R}$. Alors la variable aléatoire réelle $\alpha X + \beta$ admet une espérance et :

$$E(\alpha X + \beta) = \alpha E(X) + \beta$$

Proposition (linéarité de l'espérance) III.4.9. — Soient X et Y deux variables aléatoires réelles à densité, admettant une espérance, et $\alpha, \beta \in \mathbb{R}$.

Alors la variable aléatoire $\alpha X + \beta Y$ admet une espérance et

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

Autrement dit, l'application $X \mapsto E(X)$ est une forme linéaire de l'espace vectoriel des variables aléatoires admettant une espérance.

Théorème (de transfert) III.4.10. — Soient X une variable aléatoire réelle de densité f_X et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue par morceaux sur tout segment et telle que l'intégrale $\int_{-\infty}^{+\infty} \varphi(t) f_X(t) dt$ converge absolument.

Alors $\varphi(X)$ est une variable aléatoire réelle à densité dont l'espérance est donnée par :

$$E(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(t) f_X(t) dt$$

Définition III.4.11. — Soit $p \in \mathbb{R}, p \geq 1$. On note L^p l'ensemble des variables aléatoires réelles à densité X telles que X^p admet une espérance. On dit alors que X admet un moment d'ordre p , noté $m_p(X)$, où :

$$m_p(X) = E(X^p) = \int_{-\infty}^{+\infty} t^p f_X(t) dt.$$

Proposition III.4.12. — • Pour $p \geq 1$, L^p est un espace vectoriel.

• Soient $p, q \geq 1$ tels que $\frac{1}{p} + \frac{1}{q} = 1$ (on dit que p et q sont conjugués), $X \in L^p$ et $Y \in L^q$. Alors $XY \in L^1$ et :

$$E(|XY|) \leq (E(|X|^p))^{\frac{1}{p}} (E(|Y|^q))^{\frac{1}{q}}$$

• Soient $p, p' \geq 1, p' \leq p$. Si $X \in L^p$, alors $X \in L^{p'}$ et :

$$\left(E(|X|^{p'}) \right)^{\frac{1}{p'}} \leq \left(E(|X|^p) \right)^{\frac{1}{p}}$$

Proposition - Définition III.4.13. — Soit X une variable aléatoire réelle à densité et admettant un moment d'ordre 2.

Alors la variable aléatoire $X - E(X)$ admet un moment d'ordre 2.

La variance de X , notée $\text{Var}(X)$, est le réel

$$\text{Var}(X) = m_2(X - E(X)) = E((X - E(X))^2).$$

Proposition III.4.14. — Soit X une variable aléatoire réelle à densité et admettant un moment d'ordre 2. Alors on a $\text{Var}(X) \geq 0$.

De plus, $\text{Var}(X) = 0$ si et seulement si $X = E(X)$ presque sûrement.

Proposition III.4.15. — Soient X une variable aléatoire réelle à densité, admettant un moment d'ordre 2, et $\alpha, \beta \in \mathbb{R}$.

Alors la variable aléatoire $\alpha X + \beta$ admet une variance et

$$\text{Var}(\alpha X + \beta) = \alpha^2 \text{Var}(X).$$

Proposition (formule de Kœnig-Huygens) III.4.16. — Soit X une variable aléatoire réelle à densité et admettant un moment d'ordre 2. On a :

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

Définition III.4.17. — Soit X une variable aléatoire réelle à densité et admettant un moment d'ordre 2. L'écart-type de X , noté $\sigma(X)$, est le réel $\sigma(X) = \sqrt{\text{Var}(X)}$.

III.5 Lois usuelles à densité

III.5.1 Loi uniforme

Définition III.5.1. — Une variable aléatoire réelle X suit la loi uniforme sur $[a; b]$ si X admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a; b] \\ 0 & \text{sinon} \end{cases}$$

« X suit la loi uniforme sur $[a; b]$ » se note $X \hookrightarrow \mathcal{U}([a; b])$.

Proposition III.5.2. — La fonction de répartition F_X d'une variable aléatoire réelle X suivant la loi uniforme $\mathcal{U}([a; b])$ est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a; b] \\ 1 & \text{si } x > b \end{cases}$$

Proposition III.5.3. — Soit X une variable aléatoire réelle suivant la loi uniforme $\mathcal{U}([a; b])$. X admet une espérance et une variance égales à

$$E(X) = \frac{b+a}{2} \quad \text{et} \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

Proposition III.5.4. — Soit X une variable aléatoire réelle. On a les équivalences :

$$\begin{aligned} X \hookrightarrow \mathcal{U}([0, 1]) &\iff a + (b-a)X \hookrightarrow \mathcal{U}([a, b]) \\ X \hookrightarrow \mathcal{U}([a, b]) &\iff \frac{X-a}{b-a} \hookrightarrow \mathcal{U}([0, 1]) \end{aligned}$$

III.5.2 Loi exponentielle

Définition III.5.5. — Soit $\lambda \in \mathbb{R}_+^*$. Une variable aléatoire réelle X suit la loi exponentielle de paramètre λ si X admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

« X suit la loi exponentielle de paramètre λ » se note $X \hookrightarrow \mathcal{E}(\lambda)$.

Proposition III.5.6. — La fonction de répartition F_X d'une variable aléatoire réelle X suivant la loi exponentielle $\mathcal{E}(\lambda)$ est donnée par :

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Remarque III.5.7. — Une variable aléatoire X suivant une loi exponentielle vérifie :

$$\forall s, t \in \mathbb{R}_+^*, \quad P_{(X>s)}(X > s + t) = P(X > t)$$

Cette propriété qui traduit l'absence de mémoire de la loi exponentielle caractérise cette loi.

Proposition III.5.8. — Soit X une variable aléatoire réelle suivant la loi exponentielle $\mathcal{E}(\lambda)$. X admet une espérance et une variance égales à

$$E(X) = \frac{1}{\lambda} \quad \text{et} \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

Proposition III.5.9. — Soient X une variable aléatoire réelle et $\lambda \in \mathbb{R}_+^*$. On a les équivalences :

$$\begin{aligned} X \hookrightarrow \mathcal{E}(1) &\iff \frac{1}{\lambda}X \hookrightarrow \mathcal{E}(\lambda) \\ X \hookrightarrow \mathcal{E}(\lambda) &\iff \lambda X \hookrightarrow \mathcal{E}(1) \end{aligned}$$

III.5.3 Loi de Cauchy

Définition III.5.10. — Une variable aléatoire réelle X suit la loi de Cauchy⁵ si X admet pour densité la fonction f_X définie par

$$f_X(x) = \frac{1}{\pi(1+x^2)}$$

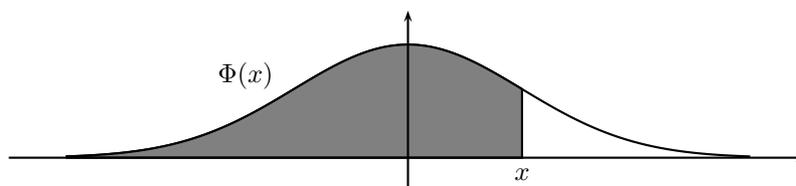
Proposition III.5.11. — Une variable aléatoire réelle X suivant la loi de Cauchy n'admet aucun moment. En particulier, X n'admet ni espérance ni variance.

III.5.4 Loi normale

Définition III.5.12. — Une variable aléatoire réelle X suit la loi normale centrée réduite si X admet pour densité la fonction φ définie par

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad \text{pour tout } x \in \mathbb{R}.$$

« X suit la loi normale centrée réduite » se note $X \hookrightarrow \mathcal{N}(0, 1)$.
On note Φ la fonction de répartition de X .

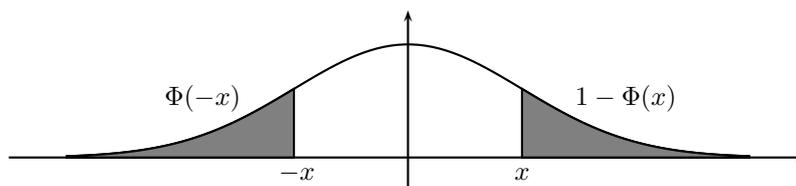


Proposition III.5.13. — Soit X une variable aléatoire réelle suivant la loi normale $\mathcal{N}(0, 1)$. X admet une espérance et une variance égales à

$$E(X) = 0 \quad \text{et} \quad \text{Var}(X) = 1.$$

Proposition III.5.14. — Pour tout $x \in \mathbb{R}$, on a :

$$\Phi(x) + \Phi(-x) = 1.$$



5. Augustin-Louis Cauchy, Paris 1789 - Sceaux 1857

Définition III.5.15. — Soient $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$. Une variable aléatoire réelle X suit la loi normale (ou de Laplace⁶-Gauss⁷) de paramètres μ et σ^2 si X admet pour densité la fonction f_X définie par

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \text{pour tout } x \in \mathbb{R}.$$

« X suit la loi normale de paramètres μ et σ^2 » se note $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

On dit également que X est une variable aléatoire gaussienne.

Proposition III.5.16. — Soit X une variable aléatoire réelle. On a les équivalences :

$$\begin{aligned} X \hookrightarrow \mathcal{N}(\mu, \sigma^2) &\Leftrightarrow X^* = \frac{X-\mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1) \\ X \hookrightarrow \mathcal{N}(0, 1) &\Leftrightarrow \sigma X + \mu \hookrightarrow \mathcal{N}(\mu, \sigma^2). \end{aligned}$$

Remarque III.5.17. — De même, si X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$, alors $aX + b$ suit la loi normale $\mathcal{N}(a\mu + b, a^2\sigma^2)$ pour tous $a, b \in \mathbb{R}$, $a \neq 0$.

Proposition III.5.18. — Soit X une variable aléatoire réelle suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$. X admet une espérance et une variance égales à

$$E(X) = \mu \quad \text{et} \quad \text{Var}(X) = \sigma^2.$$

Proposition III.5.19. — Soient X et Y deux variables aléatoires réelles indépendantes suivant les lois normales $\mathcal{N}(\mu, \sigma^2)$ et $\mathcal{N}(\mu', \sigma'^2)$ respectivement.

Alors la variable aléatoire $X + Y$ suit la loi normale $\mathcal{N}(\mu + \mu', \sigma^2 + \sigma'^2)$.

Corollaire III.5.20. — Soient X_1, \dots, X_n n variables aléatoires réelles mutuellement indépendantes telles que, pour tout $k \in \llbracket 1, n \rrbracket$, X_k suit la loi normale $\mathcal{N}(\mu_k, \sigma_k^2)$.

Alors la variable aléatoire $X_1 + \dots + X_n$ suit la loi normale $\mathcal{N}(\mu_1 + \dots + \mu_n, \sigma_1^2 + \dots + \sigma_n^2)$.

III.5.5 Loi Gamma

Définition III.5.21. — La fonction gamma, notée Γ , est définie sur \mathbb{R}_+^* par :

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt \quad \text{pour tout } x \in \mathbb{R}_+^*.$$

Proposition III.5.22. — Pour tout $x \in \mathbb{R}_+^*$, on a : $\Gamma(x+1) = x\Gamma(x)$.

Corollaire III.5.23. — Pour tout $n \in \mathbb{N}$, on a : $\Gamma(n+1) = n!$.

Définition III.5.24. — Soit $\nu \in \mathbb{R}_+^*$. Une variable aléatoire réelle X suit la loi gamma de paramètre ν si X admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(\nu)} x^{\nu-1} e^{-x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

« X suit la loi gamma de paramètre ν » se note $X \hookrightarrow \gamma(\nu)$.

Proposition III.5.25. — Soit X une variable aléatoire réelle suivant la loi gamma $\gamma(\nu)$. X admet une espérance et une variance égales à

$$E(X) = \nu \quad \text{et} \quad \text{Var}(X) = \nu$$

Proposition III.5.26. — Soient X et Y deux variables aléatoires réelles indépendantes suivant les lois gamma $\gamma(\nu)$ et $\gamma(\nu')$ respectivement.

La variable aléatoire $X + Y$ suit la loi gamma $\gamma(\nu + \nu')$.

Corollaire III.5.27. — Soient X_1, \dots, X_n n variables aléatoires réelles mutuellement indépendantes telles que, pour tout $k \in \llbracket 1, n \rrbracket$, X_k suit la loi gamma $\gamma(\nu_k)$.

Alors la variable aléatoire $X_1 + \dots + X_n$ suit la loi gamma $\gamma(\nu_1 + \dots + \nu_n)$.

6. Pierre-Simon Laplace, Beaumont-en-Auge 1749 - Paris 1827

7. Carl Friedrich Gauss, Brunswick 1777 - Göttingen 1855

Définition III.5.28. — Soit $b, \nu \in \mathbb{R}_+^*$. Une variable aléatoire réelle X suit la loi Gamma de paramètre (b, ν) si X admet pour densité la fonction f_X définie par

$$f_X(x) = \begin{cases} \frac{x^{\nu-1} e^{-\frac{x}{b}}}{b^\nu \Gamma(\nu)} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

« X suit la loi Gamma de paramètre (b, ν) » se note $X \hookrightarrow \Gamma(b, \nu)$.

Remarque III.5.29. — La loi Gamma $\Gamma(1, \nu)$ correspond à la loi gamma $\gamma(\nu)$ et la loi Gamma $\Gamma(b, 1)$ à la loi exponentielle $\mathcal{E}\left(\frac{1}{b}\right)$.

Proposition III.5.30. — Soient X une variable aléatoire réelle et $b, \nu \in \mathbb{R}_+^*$. On a les équivalences :

$$\begin{aligned} X \hookrightarrow \gamma(\nu) &\iff bX \hookrightarrow \Gamma(b, \nu) \\ X \hookrightarrow \Gamma(b, \nu) &\iff \frac{1}{b}X \hookrightarrow \gamma(\nu) \end{aligned}$$

Proposition III.5.31. — Soit X une variable aléatoire réelle suivant la loi Gamma $\Gamma(b, \nu)$. X admet une espérance et une variance égales à

$$E(X) = b\nu \quad \text{et} \quad \text{Var}(X) = b^2\nu$$

Proposition III.5.32. — Soient X et Y deux variables aléatoires réelles indépendantes suivant les lois Gamma $\Gamma(b, \nu)$ et $\Gamma(b, \nu')$ respectivement.

Alors la variable aléatoire $X + Y$ suit la loi Gamma $\Gamma(b, \nu + \nu')$.

Corollaire III.5.33. — Soient X_1, \dots, X_n n variables aléatoires réelles mutuellement indépendantes telles que X_k suit la loi Gamma $\Gamma(b, \nu_k)$.

Alors la variable aléatoire $X_1 + \dots + X_n$ suit la loi Gamma $\Gamma(b, \nu_1 + \dots + \nu_n)$.

Corollaire III.5.34. — Soient X_1, \dots, X_n n variables aléatoires réelles mutuellement indépendantes suivant la même loi exponentielle $\mathcal{E}(\lambda)$.

Alors la variable aléatoire $X_1 + \dots + X_n$ suit la loi Gamma $\Gamma\left(\frac{1}{\lambda}, n\right)$.

IV.1 Généralités

IV.1.1 Définitions

Définition IV.1.1. — Soit (Ω, \mathcal{T}) un espace probabilisable. On appelle vecteur aléatoire réel toute application mesurable de (Ω, \mathcal{T}) dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, autrement dit toute application $X = (X_1, \dots, X_n)$ de Ω dans \mathbb{R}^n telle que, pour tout $k \in \llbracket 1, n \rrbracket$, X_k soit une variable aléatoire réelle.

Proposition - Définition IV.1.2. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ à valeurs dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

L'application $\mathbb{P}_X : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1]$ définie par :

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$$

est une probabilité sur l'espace probabilisable $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Cette probabilité est appelée loi de probabilité du vecteur aléatoire X . On dit aussi que X suit la loi de probabilité \mathbb{P}_X .

Définition IV.1.3. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. Pour tout $k \in \llbracket 1, n \rrbracket$, la loi marginale de X_k est l'application $\mathbb{P}_{X_k} : \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$ définie par :

$$\mathbb{P}_{X_k}(B_k) = \mathbb{P}(X_k \in B_k) = \mathbb{P}(X^{-1}(\mathbb{R}^{k-1} \times B_k \times \mathbb{R}^{n-k}))$$

Définition IV.1.4. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire réel défini sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. La fonction de répartition de X est l'application

$$F_X : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) & \longmapsto \mathbb{P}_X(X_1 \leq x_1, \dots, X_n \leq x_n) \end{cases}$$

IV.1.2 Indépendance de variables aléatoires

Définition IV.1.5. — Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. X et Y sont indépendantes si

$$\forall A, B \in \mathcal{B}(\mathbb{R}), \quad \mathbb{P}((X \in A) \cap (Y \in B)) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Définition IV.1.6. — Soient X_1, \dots, X_n des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. X_1, \dots, X_n sont deux à deux indépendantes si, pour tous $i, j \in \llbracket 1, n \rrbracket$, $i \neq j$, X_i et X_j sont indépendantes.

Définition IV.1.7. — Soient X_1, \dots, X_n des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. X_1, \dots, X_n sont mutuellement indépendantes si

$$\forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}), \quad \mathbb{P}\left(\bigcap_{k=1}^n (X_k \in B_k)\right) = \prod_{k=1}^n \mathbb{P}(X_k \in B_k)$$

Définition IV.1.8. — Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. Les variables aléatoires $(X_n)_{n \in \mathbb{N}}$ sont mutuellement indépendantes si, pour toute partie finie non vide I de \mathbb{N} , les variables aléatoires $(X_i)_{i \in I}$ sont mutuellement indépendantes.

Théorème IV.1.9. — Soient X_1, \dots, X_n des variables aléatoires mutuellement indépendantes définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et f_1, \dots, f_n des fonctions mesurables.

Alors $f_1(X_1), \dots, f_n(X_n)$ sont des variables aléatoires mutuellement indépendantes.

Définition IV.1.10. — Soient $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$. Les variables aléatoires X_1, \dots, X_n sont mutuellement indépendantes si et seulement si

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n, \quad F_X(x_1, \dots, x_n) = \prod_{k=1}^n F_{X_k}(x_k)$$

IV.1.3 Covariance

Définition IV.1.11. — Soit (X, Y) un couple de variables aléatoires défini sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant un moment d'ordre 2.

La covariance de X et Y est le réel :

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

Si $\text{Cov}(X, Y) = 0$, X et Y sont dites non corrélées.

Proposition IV.1.12. — Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant un moment d'ordre 2.

On a :

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

Proposition IV.1.13. — Soient X, X_1, X_2, Y_1, Y_2, Y des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2, et $\alpha, \beta \in \mathbb{R}$.

On a :

- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(\alpha X_1 + \beta X_2, Y) = \alpha \text{Cov}(X_1, Y) + \beta \text{Cov}(X_2, Y)$
- $\text{Cov}(X, \alpha Y_1 + \beta Y_2) = \alpha \text{Cov}(X, Y_1) + \beta \text{Cov}(X, Y_2)$
- $\text{Cov}(X, X) = \text{Var}(X)$

Autrement dit, l'application $(X, Y) \mapsto \text{Cov}(X, Y)$ est une forme bilinéaire symétrique positive de l'espace vectoriel des variables aléatoires admettant un moment d'ordre 2.

Proposition IV.1.14. — Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant un moment d'ordre 2.

Si X et Y sont indépendantes, alors elles sont non corrélées.

Proposition IV.1.15. — Soient X_1, \dots, X_n des variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant un moment d'ordre 2.

La variable aléatoire $\sum_{k=1}^n X_k$ admet une variance égale à :

$$\text{Var} \left(\sum_{k=1}^n X_k \right) = \sum_{k=1}^n \text{Var}(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

Si, de plus, X_1, \dots, X_n sont deux à deux indépendantes, on a :

$$\text{Var} \left(\sum_{k=1}^n X_k \right) = \sum_{k=1}^n \text{Var}(X_k)$$

Définition IV.1.16. — Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant des variances non nulles.

Le coefficient de corrélation linéaire de X et Y est le réel :

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

Théorème IV.1.17. — Soient X et Y deux variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et admettant des variances non nulles.

On a :

$$|\rho(X, Y)| \leq 1$$

$|\rho(X, Y)| = 1$ si et seulement s'il existe $a, b \in \mathbb{R}$ tels que $Y = aX + b$ presque sûrement.

Définition IV.1.18. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé (Ω, \mathcal{T}, P) et tel que, pour tout $k \in \llbracket 1, n \rrbracket$, X_k admette une espérance. Le vecteur espérance de X est :

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{pmatrix}$$

Définition IV.1.19. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé (Ω, \mathcal{T}, P) et tel que, pour tout $k \in \llbracket 1, n \rrbracket$, X_k admette une variance. La matrice des covariances de X est :

$$\Gamma_X = E((X - E(X))^t (X - E(X))) = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n}$$

Proposition IV.1.20. — La matrice de covariance Γ_X est symétrique positive.

IV.2 Vecteurs aléatoires discrets

Définition IV.2.1. — Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est discret si, pour tout $k \in \llbracket 1, n \rrbracket$, X_k est une variable aléatoire discrète.

Proposition - Définition IV.2.2. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret défini sur un espace probabilisable (Ω, \mathcal{T}) . Alors la famille $((X_1 = x_1) \cap \dots \cap (X_n = x_n))_{(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)}$ est un système complet d'événements de (Ω, \mathcal{T}) appelé le système complet d'événements associé au vecteur aléatoire X .

Définition IV.2.3. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret défini sur un espace probabilisé (Ω, \mathcal{T}, P) . La loi du vecteur X ou loi conjointe de (X_1, \dots, X_n) est l'application

$$P_X : \begin{cases} X_1(\Omega) \times \dots \times X_n(\Omega) & \longrightarrow & \mathbb{R} \\ (x_1, \dots, x_n) & \longmapsto & P((X_1 = x_1) \cap \dots \cap (X_n = x_n)) \end{cases}$$

Définition IV.2.4. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret défini sur un espace probabilisé (Ω, \mathcal{T}, P) . Pour tout $k \in \llbracket 1, n \rrbracket$, la loi marginale de X_k est l'application

$$P_{X_k} : \begin{cases} X_k(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P(X_k = x) \end{cases}$$

Théorème IV.2.5. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire discret défini sur un espace probabilisé (Ω, \mathcal{T}, P) . Pour tout $k \in \llbracket 1, n \rrbracket$ et tout $x \in X_k(\Omega)$, on a :

$$P(X_k = x) = \sum_{x_1 \in X_1(\Omega)} \dots \sum_{x_{k-1} \in X_{k-1}(\Omega)} \sum_{x_{k+1} \in X_{k+1}(\Omega)} \dots \sum_{x_n \in X_n(\Omega)} P((X_1 = x_1) \cap \dots \cap (X_{k-1} = x_{k-1}) \cap (X_k = x) \cap (X_{k+1} = x_{k+1}) \cap \dots \cap (X_n = x_n))$$

Définition IV.2.6. — Soit (X, Y) un couple de variables aléatoires réelles discrètes défini sur un espace probabilisé (Ω, \mathcal{T}, P) .

Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, l'application

$$\begin{cases} X(\Omega) & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \frac{P((X = x) \cap (Y = y))}{P(Y = y)} = P_{(Y=y)}(X = x) \end{cases}$$

est appelée loi conditionnelle à $(Y = y)$ de X .

Pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, l'application

$$\begin{cases} Y(\Omega) & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \frac{P((X = x) \cap (Y = y))}{P(X = x)} = P_{(X=x)}(Y = y) \end{cases}$$

est appelée loi conditionnelle à $(X = x)$ de Y .

Définition IV.2.7. — Soit (X, Y) un couple de variables aléatoires discrètes défini sur un espace probabilisé (Ω, \mathcal{T}, P) . Les variables aléatoires X et Y sont indépendantes si

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P((X = x) \cap (Y = y)) = P(X = x)P(Y = y)$$

Définition IV.2.8. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire défini sur un espace probabilisé (Ω, \mathcal{T}, P) . Les variables aléatoires X_1, \dots, X_n sont mutuellement indépendantes si

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \quad P\left(\bigcap_{k=1}^n (X_k = x_k)\right) = \prod_{k=1}^n P(X_k = x_k)$$

Proposition IV.2.9. — Soit (X, Y) un couple de variables aléatoires discrètes défini sur un espace probabilisé (Ω, \mathcal{T}, P) et φ une application définie sur $X(\Omega) \times Y(\Omega)$. Alors l'application

$$Z : \begin{cases} \Omega & \longrightarrow & \mathbb{R} \\ \omega & \longmapsto & \varphi(X(\omega), Y(\omega)) \end{cases}$$

est une variable aléatoire réelle discrète notée $\varphi(X, Y)$.

La loi de Z est donnée par :

$$P_Z : \begin{cases} Z(\Omega) & \longrightarrow & \mathbb{R} \\ z & \longmapsto & \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ \varphi(x,y)=z}} P((X=x) \cap (Y=y)) \end{cases}$$

Théorème IV.2.10. — Soit (X, Y) un couple de variables aléatoires discrètes défini sur un espace probabilisé (Ω, \mathcal{T}, P) et φ une application définie sur $X(\Omega) \times Y(\Omega)$.

La variable aléatoire réelle discrète $\varphi(X, Y)$ admet une espérance si et seulement si la série double

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \varphi(x, y) P((X=x) \cap (Y=y))$$

est absolument convergente. En cas de convergence absolue, on a alors :

$$E(\varphi(X, Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} \varphi(x, y) P((X=x) \cap (Y=y))$$

Corollaire IV.2.11. — Soit (X, Y) un couple de variables aléatoires discrètes défini sur un espace probabilisé (Ω, \mathcal{T}, P) et admettant un moment d'ordre 2.

On a :

$$E(XY) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy P((X=x) \cap (Y=y))$$

IV.3 Vecteurs aléatoires à densité

Définition IV.3.1. — On appelle densité de probabilité sur \mathbb{R}^n toute fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ telle que $\int_{\mathbb{R}^n} f(t) dt = 1$.

Définition IV.3.2. — Soit f une densité de probabilité sur \mathbb{R}^n . Un vecteur aléatoire $X = (X_1, \dots, X_n)$ possède la loi de densité f si, pour tous intervalles I_1, \dots, I_n de \mathbb{R} , on a :

$$P((X_1 \in I_1) \cap \dots \cap (X_n \in I_n)) = \int_{I_1} \dots \int_{I_n} f(x_1, \dots, x_n) dx_1 \dots dx_n$$

Proposition IV.3.3. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de densité f_X . Alors, pour tout $k \in [1, n]$, X_k admet une densité f_{X_k} donnée par :

$$f_{X_k}(x_k) = \int_{\mathbb{R}^{n-1}} f_X(x_1, \dots, x_n) dx_1 \dots dx_{k-1} dx_{k+1} \dots dx_n$$

Proposition IV.3.4. — Soit (X, Y) un couple de variables aléatoires réelles admettant une densité $f_{(X,Y)}$ et de fonction de répartition $F_{(X,Y)}$.

Si f est continue en (x_0, y_0) , on a :

$$f(x_0, y_0) = \frac{\partial^2 F_{(X,Y)}}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 F_{(X,Y)}}{\partial y \partial x}(x_0, y_0).$$

Proposition - Définition IV.3.5. — Soit (X, Y) un couple de variables aléatoires réelles admettant une densité $f_{(X,Y)}$ et des densités marginales f_X et f_Y .

Pour tout $y \in \mathbb{R}$ tel que $f_Y(y) > 0$, la fonction définie par

$$f_{X/(Y=y)} : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \frac{f_{(X,Y)}(x, y)}{f_Y(y)} \end{cases}$$

est la densité d'une variable aléatoire, appelée densité conditionnelle de X sachant $(Y = y)$ ou encore densité de X conditionnelle à $(Y = y)$.

Pour tout $x \in \mathbb{R}$ tel que $f_X(x) > 0$, la fonction définie par

$$f_{Y/(X=x)} : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ y & \longmapsto \frac{f_{(X,Y)}(x, y)}{f_X(x)} \end{cases}$$

est la densité d'une variable aléatoire, appelée densité conditionnelle de Y sachant $(X = x)$ ou encore densité de Y conditionnelle à $(X = x)$.

Proposition IV.3.6. — Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire.

Si les variables aléatoires X_1, \dots, X_n sont mutuellement indépendantes et admettent pour densités respectives f_{X_1}, \dots, f_{X_n} , alors X admet pour densité la fonction f_X définie par :

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n, \quad f_X(x_1, \dots, x_n) = \prod_{k=1}^n f_{X_k}(x_k)$$

Réciproquement, si X admet une densité f_X , produit de n fonctions intégrables positives f_1, \dots, f_n , c'est-à-dire vérifiant

$$\forall (x_1, \dots, x_n) \in \mathbb{R}^n, \quad f_X(x_1, \dots, x_n) = \prod_{k=1}^n f_k(x_k)$$

alors les variables aléatoires X_1, \dots, X_n sont mutuellement indépendantes et, pour tout $k \in \llbracket 1, n \rrbracket$, X_k admet pour densité $f_{X_k} = \alpha_k f_k$ avec $\alpha_k > 0$ et $\prod_{k=1}^n \alpha_k = 1$.

Théorème IV.3.7. — Soit (X, Y) un couple de variables aléatoires réelles admettant une densité $f_{(X,Y)}$ et $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ un difféomorphisme.

Le vecteur aléatoire $\varphi(X, Y)$ admet une densité donnée par :

$$f_{\varphi(X,Y)}(x, y) = \begin{cases} f_{(X,Y)}(\varphi^{-1}(x, y)) |\det(J_{\varphi^{-1}}(x, y))| & \text{si } (x, y) \in \varphi(\mathbb{R}^2) \\ 0 & \text{sinon} \end{cases}$$

Théorème IV.3.8. — Soient X et Y deux variables aléatoires indépendantes admettant des densités f_X et f_Y .

Alors la variable aléatoire $X + Y$ admet une densité donnée par :

$$f_{X+Y}(x) = \int_{-\infty}^{+\infty} f_X(t) f_Y(x-t) dt = \int_{-\infty}^{+\infty} f_Y(t) f_X(x-t) dt$$

Proposition IV.3.9. — Soit (X, Y) un couple de variables aléatoires admettant une densité $f_{(X,Y)}$ et φ une application définie sur $(X, Y)(\Omega)$.

La variable aléatoire $\varphi(X, Y)$ admet une espérance si et seulement si l'intégrale $\int_{\mathbb{R}^2} \varphi(x, y) f_{(X,Y)}(x, y) dx dy$ est absolument convergente.

En cas de convergence absolue, on a alors :

$$E(\varphi(X, Y)) = \int_{\mathbb{R}^2} \varphi(x, y) f_{(X,Y)}(x, y) dx dy$$

Corollaire IV.3.10. — Soit (X, Y) un couple de variables aléatoires admettant une densité $f_{(X,Y)}$ et un moment d'ordre 2.

On a :

$$E(XY) = \int_{\mathbb{R}^2} xy f_{(X,Y)}(x, y) dx dy$$

IV.4 Vecteurs gaussiens

Définition IV.4.1. — Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est dit gaussien si toute combinaison linéaire de ses composantes, soit $\sum_{k=1}^n \alpha_k X_k$ avec $\alpha_1, \dots, \alpha_n \in \mathbb{R}$, est une variable aléatoire gaussienne.

Théorème IV.4.2. — Soit X un vecteur aléatoire gaussien de \mathbb{R}^n , non dégénéré, c'est-à-dire dont la matrice de covariance Γ_X est inversible. X admet pour densité de probabilité la fonction f_X définie par :

$$\forall x \in \mathbb{R}^n, \quad f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{\det(\Gamma_X)}} \exp\left(-\frac{1}{2}{}^t(x - \mathbb{E}(X))\Gamma_X^{-1}(x - \mathbb{E}(X))\right).$$

De plus, $Y = A^{-1}(X - \mathbb{E}(X))$, où A est une matrice carrée inversible telle que $\Gamma_X = A^t A$, est un vecteur gaussien dont les composantes sont des variables gaussiennes centrées réduites mutuellement indépendantes.

Théorème IV.4.3. — Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien de matrice de covariance Γ_X . Les variables aléatoires X_1, \dots, X_n sont deux à deux non corrélées (i.e. Γ_X est diagonale) si et seulement si elles sont mutuellement indépendantes.

V.1 Quelques résultats

V.1.1 Inégalités de Markov et de Bienaymé-Tchebychev

Théorème (inégalité de Markov¹) V.1.1. — Soit X une variable aléatoire réelle positive définie sur un espace probabilisé (Ω, \mathcal{T}, P) et admettant une espérance.

Pour tout $a \in \mathbb{R}_+^*$, on a :

$$P(X \geq a) \leq \frac{E(X)}{a}$$

Proposition (inégalité de Bienaymé²-Tchebychev³) V.1.2. — Soit X une variable aléatoire réelle positive définie sur un espace probabilisé (Ω, \mathcal{T}, P) et admettant une variance.

Pour tout $a \in \mathbb{R}_+^*$, on a :

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

V.1.2 Lemme de Borel-Cantelli

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de Ω .

Pour tout $n \in \mathbb{N}$, la suite $\left(\bigcup_{k \geq n} A_k \right)_{n \in \mathbb{N}}$ est une suite décroissante (au sens de l'inclusion) d'événements de Ω .

Par définition, sa limite est $\bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k$. Cette limite s'appelle la limite supérieure de la suite $(A_n)_{n \in \mathbb{N}}$:

$$\limsup A_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k$$

Pour tout $n \in \mathbb{N}$, la suite $\left(\bigcap_{k \geq n} A_k \right)_{n \in \mathbb{N}}$ est une suite croissante (au sens de l'inclusion) d'événements de Ω . Par définition, sa limite est $\bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k$. Cette limite s'appelle la limite inférieure de la suite $(A_n)_{n \in \mathbb{N}}$:

$$\liminf A_n = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k$$

L'ensemble $\limsup A_n$ est l'ensemble des $\omega \in \Omega$ qui appartiennent à une infinité de A_n . Ainsi, l'événement $\limsup A_n$ signifie « les événements A_n sont réalisés une infinité de fois » ce que l'on abrège par « A_n infiniment souvent » ou encore « A_n i.s. ».

L'ensemble $\liminf A_n$ est l'ensemble des $\omega \in \Omega$ qui, à partir d'un certain rang $n(\omega)$, appartiennent à tous les A_n .

1. Andreï Markov, Saint Pétersbourg 1903 - Moscou 1979

2. Jules Bienaymé, Paris 1796 - Paris 1878

3. Pafnouti Tchebychev, Okatovo 1821 - Saint Pétersbourg 1894

On a l'inclusion :

$$\liminf A_n \subset \limsup A_n$$

ainsi que les égalités :

$$\overline{\limsup A_n} = \liminf \overline{A_n} \quad \text{et} \quad \overline{\liminf A_n} = \limsup \overline{A_n}$$

Théorème (lemme de Borel⁴-Cantelli⁵) V.1.3. — Soient $(A_n)_{n \in \mathbb{N}}$ une suite d'événements sur un espace probabilisé (Ω, \mathcal{T}, P) .

(i) Si la série de terme général $P(A_n)$ converge, alors $P(\limsup A_n) = 0$.

(ii) Si les événements A_n sont mutuellement indépendants et si la série de terme général $P(A_n)$ diverge, alors $P(\limsup A_n) = 1$.

V.2 Convergence des variables aléatoires

Les variables aléatoires considérées sont définies sur le même espace probabilisé (Ω, \mathcal{T}, P) .

V.2.1 Différents modes de convergence

Définition (convergence en probabilité) V.2.1. — Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires converge en probabilité vers une variable aléatoire X si, pour tout $\varepsilon > 0$, on a :

$$\lim_{n \rightarrow +\infty} P(|X_n - X| \geq \varepsilon) = 0$$

On note : $X_n \xrightarrow{P} X$.

Définition (convergence presque sûre) V.2.2. — Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires converge presque sûrement vers une variable aléatoire X si on a :

$$P\left(\left\{\omega \in \Omega; \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

On note : $X_n \xrightarrow{p.s.} X$

Proposition V.2.3. — Une suite $(X_n)_{n \in \mathbb{N}}$ converge presque sûrement vers X si et seulement si

$$\forall \varepsilon > 0, \quad P(\limsup (|X_n - X| \geq \varepsilon)) = 0$$

ou encore

$$\forall \varepsilon > 0, \quad \lim_{m \rightarrow +\infty} P(\sup_{n \geq m} |X_n - X| \geq \varepsilon) = 0$$

Définition (convergence en moyenne d'ordre p) V.2.4. — Soit $p \geq 1$. Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires de $L^p(\Omega, \mathcal{T}, P)$ converge dans L^p (ou en moyenne d'ordre p) vers une variable aléatoire X si $X \in L^p(\Omega, \mathcal{T}, P)$ et si

$$\lim_{n \rightarrow +\infty} E(|X_n - X|^p) = 0.$$

On note : $X_n \xrightarrow{L^p} X$.

Définition (convergence en loi) V.2.5. — Une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires converge en loi vers une variable aléatoire X si $\lim_{n \rightarrow +\infty} F_{X_n}(t) = F_X(t)$ en tout point de continuité t de F_X .

On note $X_n \xrightarrow{\mathcal{L}} X$.

Théorème V.2.6. — Les assertions suivantes sont équivalentes :

(i) la suite $(X_n)_{n \in \mathbb{N}}$ converge en loi vers X ;

(ii) pour toute fonction continue bornée f , on a : $\lim_{n \rightarrow +\infty} \int_{\Omega} f(X_n) dP = \int_{\Omega} f(X) dP$.

V.2.2 Comparaison des différents modes de convergence

$$X_n \xrightarrow{p.s.} X \implies X_n \xrightarrow{P} X \implies X_n \xrightarrow{\mathcal{L}} X$$

$$\uparrow$$

$$X_n \xrightarrow{L^p} X$$

4. Émile Borel, Saint-Afrique 1871 - Paris 1956

5. Francesco Cantelli, Palerme 1875 - Rome 1966

V.3 Lois des grands nombres

Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires, on pose $S_n = \sum_{k=1}^n X_k$.

V.3.1 Loi faible des grands nombres

Théorème (loi faible des grands nombres) V.3.1. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, admettant un moment d'ordre 2 et deux à deux non corrélées. On suppose que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \mathbb{E}(X_k) = m \quad \text{et} \quad \lim_{n \rightarrow +\infty} \frac{1}{n^2} \sum_{k=1}^n \text{Var}(X_k) = 0$$

Alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge en probabilité vers m .

Corollaire (loi faible des grands nombres) V.3.2. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, identiquement distribuées, deux à deux non corrélées et admettant un moment d'ordre 2.

Alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge en probabilité vers $\mathbb{E}(X_1)$.

V.3.2 Loi forte des grands nombres

Théorème (de Kolmogorov⁶, loi forte des grands nombres) V.3.3. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et mutuellement indépendantes. On suppose que :

$$\lim_{n \rightarrow +\infty} \mathbb{E}(X_n) = m \quad \text{et} \quad \sum_{k=1}^{+\infty} \frac{\text{Var}(X_k)}{k^2} < +\infty$$

Alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge presque sûrement vers m .

Remarque V.3.4. — Sous les hypothèses précédentes, $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge également dans L^2 vers m .

Théorème (de Kolmogorov, loi forte des grands nombres) V.3.5. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, mutuellement indépendantes et identiquement distribuées.

Si $\mathbb{E}(|X_1|) < +\infty$, alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.

La démonstration des deux théorèmes précédents repose sur des lemmes techniques.

En revanche, on pourra s'intéresser aux deux résultats suivants dont la preuve est abordable.

Proposition V.3.6. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, mutuellement indépendantes, identiquement distribuées et admettant un moment d'ordre 4.

Alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.

Théorème V.3.7. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, mutuellement indépendantes, identiquement distribuées et admettant un moment d'ordre 2.

Alors la suite de variables aléatoires $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge presque sûrement vers $\mathbb{E}(X_1)$.

Remarque V.3.8. — Sous les hypothèses précédentes, $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}^*}$ converge également dans L^2 vers $\mathbb{E}(X_1)$.

V.4 Approximations

V.4.1 Approximation de la loi binomiale par la loi de Poisson

Proposition V.4.1. — Soient $\lambda \in \mathbb{R}_+^*$ et $(p_n)_{n \in \mathbb{N}^*}$ une suite de réels de $]0; 1[$ tels que $\lim_{n \rightarrow +\infty} np_n = \lambda$.

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires telle que X_n suit la loi binomiale $\mathcal{B}(n, p_n)$.

Alors $(X_n)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire X qui suit la loi de Poisson $\mathcal{P}(\lambda)$.

6. Andreï Kolmogorov, Tambov 1903 - Moscou 1987

Dans la pratique, la loi binomiale $\mathcal{B}(n, p)$ peut être approchée par la loi de Poisson $\mathcal{P}(np)$ si

$$n \geq 30, \quad p \leq 0,1, \quad \text{et} \quad np < 15.$$

V.4.2 Approximation de la loi binomiale par la loi normale

Théorème (De Moivre⁷-Laplace) V.4.2. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires telle que X_n suit la loi binomiale $\mathcal{B}(n, p)$.

Alors $\left(\frac{X_n - np}{\sqrt{np(1-p)}} \right)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Dans la pratique, la loi binomiale $\mathcal{B}(n, p)$ peut être approchée par la loi normale $\mathcal{N}(np, np(1-p))$ si

$$n \geq 30, \quad np \geq 15 \quad \text{et} \quad np(1-p) > 5.$$

V.5 Théorème limite central

Si $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables aléatoires, on pose $S_n = \sum_{k=1}^n X_k$.

Théorème V.5.1. — Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires non constantes, mutuellement indépendantes, identiquement distribuées et admettant un moment d'ordre 2.

Alors $\left(\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\text{Var}(S_n)}} \right)_{n \in \mathbb{N}^*}$ converge en loi vers une variable aléatoire qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$.

7. Abraham De Moivre, Vitry-le-François 1667 - Londres 1754

BIBLIOGRAPHIE

- [1] P. Barbe, M. Ledoux (2007). *Probabilité*, EDP Sciences, Les Ulis
- [2] C. Berge (1968). *Principes de combinatoire*, Dunod, Paris
- [3] P. Brémaud (deuxième édition 2009), *Initiation aux probabilités et aux chaînes de Markov*, Springer-Verlag, Berlin
- [4] D. Foata, A. Fuchs (1998). *Calcul des probabilités*, Dunod, Paris
- [5] G. Grimmett, D. Stirzaker (third edition 2001). *Probability and random processes*, Oxford university press, Oxford
- [6] J. Jacod, P. Protter (2003). *L'essentiel en théorie des probabilités*, Cassini, Paris
- [7] J.-Y. Oувrard (1998). *Probabilités tome 1. CAPES - Agrégation*, Cassini, Paris
- [8] J.-Y. Oувrard (1998). *Probabilités tome 2. Master -Agrégation*, Cassini, Paris
- [9] S.M. Ross (1987). *Initiation aux probabilités*, Presses polytechniques romandes, Lausanne
- [10] J. Stoyanov (1989). *Counterexamples in probability*, Wiley, New-York