République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Abou Bekr Belkaid Tlemcen

Faculté des Sciences

Probabilités

Boukhari Fakhreddine

Département de Mathématiques
f_boukhari@yahoo.fr
Faculté des Sciences
BP 119 Tlemcen

" Comme un aveugle n'a aucune idée des couleurs, de même nous n'avons aucune idée de la manière dont Dieu infiniment sage perçoit et comprend toutes choses."

Isaac Newton.

Table des matières

Introduction

Ce manuscrit est une introduction au calcul des probabilités, il est destiné aux étudiants de la deuxième année de la filière Génie Productique. Il peut aussi être utilisé par les étudiants de la deuxième année de la filière MI option informatique.

Ce polycopié ne peut être considéré comme un ouvrage de référence, il est écrit dans le but de servir comme aide mémoire pour un étudiant abordant pour la première fois le cours de probabilités.

Il est constitué de six chapitres

Le premier chapitre est un rappel des outils de base de l'analyse combinatoire, après un bref survol de la théorie des ensembles et des fonctions, on y expose les principes fondamentaux de l'analyse combinatoire.

Dans le second chapitre on aborde la théorie moderne du calcul des probabilités en donnant la définition mathématique d'un espace de probabilités, nous avons essayé de faire appel le moins possible aux notions de la théorie de la mesure, qui reste malgré tout essentielle pour une approche rigoureuse du calcul des probabilités. Nous avons introduit la notion de probabilité conditionnelle ainsi que la notion d'indépendance pour les événements qui reste une notion propre à la théorie de la probabilité.

Le troisième chapitre est consacré aux variables aléatoires, après la définition de cette notion nous étudions en détail les deux grandes familles de variables aléatoire à savoir les variables discrètes et les variables absolument continues, nous définissons ensuite la fonction génératrice des moments ainsi que la fonction caractéristiques essentielle pour l'identification des lois ainsi que pour le problème de convergence en loi.

Dans le quatrième chapitre on donne les principales lois de probabilités, on y traite aussi le problème de changement de variable ainsi qu'une première approche concernant la convergence en loi d'une loi binomiale vers une loi de Poisson.

Le cinquième chapitre constitue une introduction aux vecteurs aléatoires. Après avoir enseigné le cours de probabilités aux étudiants de la deuxième année informatique à plusieurs reprises, j'ai pu constater amèrement la difficulté d'aborder ce sujet dans sa généralité au vu des difficultés que rencontre de nombreux étudiants à manipuler les intégrales multiples, j'ai dû donc me restreindre aux couples aléatoires, en espérant des jours meilleurs.

En lisant le dernier chapitre, des collègues ayant enseigné ce cours me traiteront peut être d'optimiste, ce qui n'est pas mon cas malheureusement, en effet ce chapitre est consacré à la délicate question de la convergence en théorie des probabilités est ses corollaires légendaires : les lois des grands nombres et le théorème de la limite centrale. Elle est délicate car il y'a au moins quatre modes de convergence, d'autre part on est amené dans au moins un mode de convergence de traiter avec le fameux *epsilon* qui fait tant peur à nos étudiants (pourtant il a l'air sympathique), ajoutant à cela le fait que dans le plus important mode de convergence (la convergence en loi), on est amené à traiter non pas avec la suite de variables aléatoires étudiées mais avec sa suite de lois, enfin pour couronner le tout, comme en analyse, montrer la convergence ou la divergence revient souvent à faire des majorations ou des minorations, ce que beaucoup d'étudiants rechignent à faire si on n'insiste pas.

A la fin de chaque chapitre nous proposons une série d'exercices à difficulté variable, afin d'aider l'étudiant studieux à mieux assimiler le contenu de ce cours.

Cet ouvrage ne prétend à aucune originalité, le contenu exposé est standard, il fait partie de la plupart des livres traitant de la théorie moderne des probabilités. Comme chaque travail académique, il contient sûrement des erreurs et je remercie par avance tout collègue et tout étudiant qui me fera part de ses remarques et critiques.

Je remercie mon collègue le Professeur Abdellaoui Boumédiène pour son encouragement, son soutient et son enthousiasme. J'exprime ma gratitude et mon profond respect à mon collègue Miri Sofiane Maitre de conférences à l'université de Tlemcen, pour ses qualités humaines et pour avoir lu et corrigé ce manuscrit.

Chapitre 1

Analyse combinatoire

1.1 Rappels et Compléments

Dans tout ce qui suit on désigne par Ω un ensemble non-vide.

1.1.1 Théorie des ensembles

- L'ensemble vide est l'ensemble qui ne contient aucun élément, on le note $\varnothing$.
- $\mathcal{P}(\Omega)$, désigne l'ensemble des toutes les parties de Ω.
- $A \in \mathcal{P}(\Omega) \Leftrightarrow A \subset \Omega$.

Exemple 1.1. Soit $\Omega = \{a, b, c\}$, alors $\mathcal{P}(\Omega) = \{\varnothing, \Omega, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}\}$.

Soient à présent $A, B \in \mathcal{P}(\Omega)$, alors

- $\overline{A} = C_A = \{\omega \in \Omega;\ \omega \notin A\}$.
- $A \cup B = \{\omega \in \Omega;\ \omega \in A \text{ ou } \omega \in B\}$.
- $A \cap B = \{\omega \in \Omega;\ \omega \in A \text{ et } \omega \in B\}$.
- $A \backslash B = A \cap \overline{B} = \{\omega \in \Omega;\ \omega \in A \text{ et } \omega \notin B\}$.
- $A \Delta B = A \backslash B \ \cup\ B \backslash A = B \Delta A$.
- $A \subset B \Leftrightarrow \Big[\omega \in A \Rightarrow \omega \in B\Big]$.
- $A = B \Leftrightarrow \Big[A \subset B \text{ et } B \subset A\Big]$.

Proposition 1.1. Soient A, B et C trois sous ensembles de Ω. On a

1/ $A \cap A = A$

2/ $A \cup A = A$

3/ $A \cap B = B \cap A$ (*commutativité de l'intersection*)

4/ $A \cup B = B \cup A$ (*commutativité de l'union*)

5/ $(A \cap B) \cap C = A \cap (B \cap C)$ (*associativité de l'intersection*)

6/ $(A \cup B) \cup C = A \cup (B \cup C)$ (*associativité de l'union*)

7/ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (*Distributivité de l'union sur l'intersection*)

8/ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (*Distributivité de l'intersection sur l'union*)

9/ $\overline{A \cap B} = \overline{A} \cup \overline{B}$

10/ $\overline{A \cup B} = \overline{A} \cap \overline{B}$

11/ $\overline{A} \cap A = \varnothing$

12/ $\overline{A} \cup A = \Omega$

On généralise les notions d'intersection et d'union à une infinité d'ensembles de la manière suivante :

Définition 1.1. Soient I un ensemble non-vide, $\{A_i,\ i \in I\}$ une famille d'éléments de $\mathcal{P}(\Omega)$, alors

– $\bigcup_{i \in I} A_i = \{\omega \in \Omega / \ \exists i \in I \text{ tel que } \ \omega \in A_i\}$.

– $\bigcap_{i \in I} A_i = \{\omega \in \Omega / \ \omega \in A_i,\ \forall i \in I\}$.

Ainsi

– $\omega \in \bigcup_{i \in I} A_i \iff \exists i \in I, \omega \in A_i$.

– $\omega \in \bigcap_{i \in I} A_i \iff \omega \in A_i,\ \forall i \in I$.

Définition 1.2. Soit $A \in \mathcal{P}(\Omega)$, on note par $card(A)$, le nombre (éventuellement infini) d'éléments de A. On dit que A est en ensemnle fini si $card(A)$ est un nombre fini.

Exemple 1.2. $card(\emptyset) = 0,\ card(\{a, b, c, d\}) = 4,\ card(\mathbb{N}) = +\infty$.

Dans le cas d'un ensemble fini, il n'est pas difficile de montrer le résultat suivant :

Proposition 1.2. Soient Ω un ensemble fini, $A, B \in \mathcal{P}(\Omega)$. Alors

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

En particulier, si $A \cap B = \emptyset$, alors

$$card(A \cup B) = card(A) + card(B).$$

Définition 1.3. Soient Ω_1 et Ω_2 deux ensembles non-vides, le produit cartésien de Ω_1 et Ω_2 est l'ensemble noté $\Omega_1 \times \Omega_2$ défini par

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) \,;\ \omega_1 \in \Omega_1 \text{ et } \omega_2 \in \Omega_2 \}$$

Remarque 1.1. $\Omega_1 \times \Omega_2$ est l'ensemble des couples (ω_1, ω_2) où $\omega_1 \in \Omega_1$ et $\omega_2 \in \Omega_2$. Il faut noter que (ω_1, ω_2) est appelé couple ordonné, ainsi il y a une différence entre (ω_1, ω_2) et (ω_2, ω_1).

$$[(\omega_1, \omega_2) = (\omega_1', \omega_2')] \Leftrightarrow [\omega_1 = \omega_1' \text{ et } \omega_2 = \omega_2']$$

De la même manière on définit le produit cartésien de plusieurs ensembles :

Définition 1.4. Soient $k \geq 2$, $\{\Omega_n,\ n \geq 1\}$, une familles infinie d'ensembles non-vides.

- On appelle k-uplet une disposition ordonnée de k éléments de type $(\omega_1, \omega_2, , ..., \omega_k)$ où $\omega_i \in \Omega_i$ pour chaque $1 \leq i \leq k$.
- Le produit cartésien $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_k$ est l'ensemble des k-uplets, i. e.

$$\Omega_1 \times \Omega_2 \times \cdots \times \Omega_k = \{(\omega_1, \omega_2, , ..., \omega_k),\ \omega_i \in \Omega_i, \ \text{ pour } 1 \leq i \leq k\}.$$

1.1.2 Applications

Définition 1.5. On appelle application d'un ensemble E vers un ensemble F; toute correspondance X qui associe à tout élément $\omega \in E$ un et un seul élément

$y = X(\omega) \in F$. On note

$$\begin{aligned} X: \quad & E \longrightarrow F \\ & \omega \longmapsto X(\omega) \end{aligned}$$

Remarque 1.2. On rappelle que :

- E est appelé l'ensemble de départ.
- F est appelé l'ensemble d'arrivée.
- $y = X(\omega)$ est appelé l'image de ω par X.
- ω est appelé un antécédent de $y = X(\omega)$.

Définition 1.6. Soient $X : E \longrightarrow F$ une application, $A \subset E$ et $B \subset F$.

1. Image directe : L'image directe de A par X notée $X(A)$, est le sous ensemble de F défini par

$$X(A) = \{y \in F;\ \exists\ \omega \in A \text{ tel que } y = X(\omega)\}$$

2. Image réciproque : L'image réciproque de B par X notée $X^{-1}(B)$, est le sous ensemble de E défini par

$$X^{-1}(B) = \{\omega \in E;\ X(\omega) \in B\ \}$$

Remarque 1.3. Il faut noter que

1. l'image d'un ensemble A, est tout simplement le sous ensemble de F constitué des images de tous les éléments de A.
2. l'image réciproque de B est le sous ensemble de E constitué des éléments dont l'image est dans B.
3. pour parler de l'image réciproque d'un ensemble par une application X, il n'est pas nécessaire d'exiger que X soit bijective contrairement à l'image réciproque d'un élément par une application X qui n'a de sens que si X est bijective.

Proposition 1.3. Soient $X : E \longrightarrow F$ une application, $A \subset E$, $B \subset E$ et $C \subset F$, $D \subset F$. On a

1. $A \subset B \Rightarrow X(A) \subset X(B)$

2. $X(A \cup B) = X(A) \cup X(B)$
3. $X(A \cap B) \subset X(A) \cap X(B)$
4. Si X est injective alors, $X(A \cap B) = X(A) \cap X(B)$
5. $C \subset D \Rightarrow X^{-1}(C) \subset X^{-1}(D)$
6. $X^{-1}(C \cup D) = X^{-1}(C) \cup X^{-1}(D)$
7. $X^{-1}(C \cap D) = X^{-1}(C) \cap X^{-1}(D)$

Nous terminons ce paragraphe par introduire la fonction indictarice d'un ensemble, très importante en théorie des probabilités :

Définition 1.7. Soient Ω un ensemble non vide, $A \in \mathcal{P}(\Omega)$. La fonction indicatrice de A, notée $\mathbb{1}_A$ est définie par $\mathbb{1}_A : \ \Omega \longrightarrow \{0,1\}$, avec

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in \overline{A} \end{cases}$$

Cette fonction possède les propriétés suivantes :

Proposition 1.4. Soient $A, B \in \mathcal{P}(\Omega)$, alors

- $\mathbb{1}_\Omega = 1$ et $\mathbb{1}_\emptyset = 0$.
- $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B}$.
- $\mathbb{1}_{A \cup B} = \max(\mathbb{1}_A, \mathbb{1}_B)$.
- $\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$
- $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B = \min(\mathbb{1}_A, \mathbb{1}_B)$

1.2 Eléments d'analyse combinatoire

Supposons à présent que $card(\Omega) = n$, où $n \geq 1$.

Définition 1.8. Pour tout $i \geq 1$, on définit le factoriel de i par $i! = 1 \times 2 \times \cdots \times i$. On conviendra que $0! = 1$.

Remarque 1.4. Il est facile de voir que pour tout $i \geq 1$,

$$i! = (i-1)! \times i = (i-2)! \times (i-1) \times i = \cdots .$$

Définition 1.9. On appelle arrangement avec répétition de k éléments choisis parmi n, un k-uplet $(\omega_1, \omega_2, , ..., \omega_k)$ où $\omega_i \in \Omega$ pour tout $1 \leq i \leq k$.

Exemple 1.3. Soient $\Omega = \{a, b, c\}$, $k = 2$, les arrangement avec répétition de 2 éléments choisis dans Ω sont : $(a, a), (b, b), (c, c), (a, b), (b, a), (a, c), (c, a), (b, c), (c, b)$.

Remarque 1.5. Il faut bien noter que dans un arrangement avec répétition, un éléments peut figurer plusieurs fois, et que l'ordre des éléments est pris en considération, ainsi dans l'exemple précédents les éléments (a, b) et (b, a) sont différents, ils sont donc comptés tous les deux.

Proposition 1.5. Le nombre d'arrangements avec répétition de k éléments choisis parmi n est n^k.

Exemple 1.4. Pour utiliser une carte de retrait bancaire, on a besoin d'un code de 4 chiffres. Combien de codes peut-on générer ?

SOLUTION:

Chaque code est un nombre de 4 chiffres de type $abcd$, où $a, b, c, d \in \{0, 1, ..., 9\}$. Ainsi le nombre de codes possibles est $10 \times 10 \times 10 \times 10 = 10^4$ codes.

Définition 1.10. Soient $1 \leq k \leq n$, on appelle arrangement sans répétition de k éléments parmi n une disposition ordonnée sans répétition de k éléments choisis dans un ensemble de cardinal n.

Exemple 1.5. Soient $\Omega = \{a, b, c\}$, $k = 2$, les arrangements sans répétition de 2 éléments choisis dans Ω sont : $(a, b), (b, a), (a, c), (c, a), (b, c), (c, b)$.

Remarque 1.6. Dans un arrangements sans répétition, on a pas le droit de choisir un élément plus qu'une fois dans une disposition, c'est pour cette raison que l'élément (a, a) ne figure pas dans l'exemple précédent, cependant l'ordre des éléments est pris en considération.

Proposition 1.6. Le nombre d'arrangements sans répétition de k éléments parmi n est noté A_n^k, il est donné par la formule :

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

Exemple 1.6. On veut former un mot de 5 lettres avec l'alphabet français avec la condition que chaque lettre apparaît au plus une fois (*On ne tient pas compte du sens du mot*).

SOLUTION:

Puisque on exige qu'une lettre ne peut appraître qu'au plus une fois, il ne peut y avoir de répétition, c'est donc un arrangement sans répétition, ainsi on peut former A_{26}^{5} mots.

Définition 1.11. Soient $n \geq 1$, Ω un ensemble avec $card(\Omega) = n$. On appelle permutation sans répétition de n éléments une disposition ordonnée des éléments de Ω où chaque éléments figure une seule fois et une seule.

Exemple 1.7. Soit $\Omega = \{a, b, c\}$, les permutations sans répétition des Ω sont : *abc bac bca cba cab acb*.

Proposition 1.7. Le nombre de permutations sans répétition d'éléments d'un ensemble de cardinal n est $n!$

Exemple 1.8. De combien de façons peut-on ranger 4 livres différents sur une étagère ?

SOLUTION:

Il y a 4! manières de ranger 4 livres différents sur une étagère.

Définition 1.12. Soient $k \geq 1$, $\Omega_1, \Omega_2, ..., \Omega_k$ des ensembles différents contenant respectivement $n_1, n_2, ..., n_k$ éléments identiques. Posons $n = n_1 + n_2 + ... + n_k$. On appelle permutation avec répétition de n éléments une disposition ordonnée des ces éléments, où chaque éléments figure une seule fois et une seule.

Exemple 1.9. Considérons la disposition : *aabcc*, elle est composée de 3 groupes distincts, contenant respectivement 2, 1 et 2 éléments. Les permutations possibles de cette disposition sont : *aabcc*, *aacbc*, *aaccb*, *acacb*, *accab*, *acabc*,...etc

Proposition 1.8. Le nombre de permutations avec répétition d'un groupe de n éléments composé de k sous groupes différents, contenant respectivement $n_1, n_2, ...,$ n_k éléments identiques est égale à

$$\frac{n!}{n_1!\ n_2! \cdots\ n_k!}$$

Exemple 1.10. Le nombre de permutaions des éléments : *aabcc* est donc $\frac{5!}{2!1!2!} = 30$, alors que le nombre de permutations (sans répétition) des éléments : *abcde* est $5! = 120$. Cette différence est dûe au fait qu'on ne comptabilise les permutations effectuées au sein d'un même groupe qu'une seule fois.

Définition 1.13. Soient $1 \leq k \leq n$, Ω un ensemble contenant n éléments. On appelle combinaison sans répétition de k éléments de Ω une disposition non ordonnée et sans répétition de k éléments de Ω.

Remarque 1.7. D'après cette définition, on remarque que :

– Une combinaison sans répétition de k éléments parmi n est tout simplement un ensemble contenant k éléments qu'on choisis dans un plus grand ensemble de cardinal n.
– La différence entre une combinaison sans répétition et un arrangment sans répétition est donc le fait de considerer que l'ordre n'est pas important pour la première, alors qu'il est important pour un arrangement, ceci implique que le nombre de combinaisons sans répétition sera plus petit que le nombre d'arrangments sans répétition.

Exemple 1.11. Soit $\Omega = \{a, b, c\}$, les combinaisons sans répétition de 2 éléments de Ω sont : $\{a, b\}$, $\{a, c\}$ et $\{c, b\}$.

Proposition 1.9. Soient $0 \leq k \leq n$. Le nombre de combinaisons sans répétition de k éléments parmi n est noté C_n^k, il est donné par la formule :

$$C_n^k = \frac{n!}{k!\ (n-k)!}.$$

Exemple 1.12. Une urne contient 9 boules : 4 blanches, 2 noires et 3 rouges. On tire 3 boules de cette urne, de combien de manières peut-on tirer :

– Deux boules blanches et une noire.
– Une boule de chaque couleur.

Solution:

– On a C_4^2 façons de tirer deux boules blanches et C_2^1 façons de tirer une boule noire, ainsi on a $C_4^2 \times C_2^1 = 12$ manières de tirer deux boules blanches et une noire.

- On a C_4^1 façons de tirer une boule blanche, C_2^1 façons de tirer une boule noire et C_3^1 façons de tirer une boule rouge, ainsi on a $C_4^1 \times C_2^1 \times C_3^1 = 24$ manières de tirer une boule de chaque couleur.

Proposition 1.10. Pour $1 \leq k \leq n$, C_n^k est le nombre des sous ensembles contenant k éléments qu'on peut former à partir d'un ensemble de n éléments.

Remarque 1.8. Soient $0 \leq k \leq n$.

- Il n'est pas difficile de voir que $A_n^k = k!C_n^k$, ceci est dû au fait que chaque ensemble contenant k éléments genère $k!$ dispositions (permutations) si on considère que l'ordre est important ; et une seule disposition dans le cas où l'ordre n'est pas pris en compte.
- $C_n^0 = 1$, puisque C_n^0 est le nombre de sous ensembles ne contenant aucun éléments dans un ensemble à n éléments, c'est donc l'ensemble vide et il n'y en a qu'un seul.
- $C_n^n = 1$, puisque C_n^n est le nombre de sous ensembles contenant n éléments dans un ensemble à n éléments, c'est donc l'ensemble tout entier et il n'y en a qu'un seul.
- $C_n^1 = n$, puisque C_n^1 est le nombre de sous ensembles ne contenant qu'un seul élément (singletons) dans un ensemble à n éléments, il y a donc n singletons.

1.3 Exercices

Exercice 1.1. Soient $X : E \longrightarrow F$ une application, $A \subset E$, $B \subset E$ et $C \subset F$, $D \subset F$. Montrer que

1. $A \subset B \Rightarrow X(A) \subset X(B)$
2. $X(A \cup B) = X(A) \cup X(B)$
3. $X(A \cap B) \subset X(A) \cap X(B)$
4. Si X est <u>injective</u> alors, $X(A \cap B) = X(A) \cap X(B)$
5. $C \subset D \Rightarrow X^{-1}(C) \subset X^{-1}(D)$
6. $X^{-1}(C \cup D) = X^{-1}(C) \cup X^{-1}(D)$
7. $X^{-1}(C \cap D) = X^{-1}(C) \cap X^{-1}(D)$

Exercice 1.2. Soient Ω, F deux ensembles non-vides, $X : \Omega \to F$ une application, $A, B \in \mathcal{P}(\Omega)$ et $C, D \in \mathcal{P}(F)$.

1. Montrer que

$$X^{-1}(C \cap D) = X^{-1}(C) \cap X^{-1}(D)$$

$$X^{-1}(C \cup D) = X^{-1}(C) \cup X^{-1}(D)$$

$$X^{-1}(\overline{C}) = \overline{X^{-1}(C)}$$

$$X(A \cup B) = X(A) \cup X(B)$$

$$X(A \cap B) \subset X(A) \cap X(B).$$

2. Montrer qu'en général, la dernière inclusion est stricte.
3. Montrer que si X est injective alors la dernière inclusion est une égalité.

Exercice 1.3. Soient $n \in \mathbb{N}$, Ω un ensemble fini avec $card(\Omega) = n$. Montrer par récurrence que $card(\mathcal{P}(\Omega)) = 2^n$.

Exercice 1.4. Soient $1 \leq k \leq n$. Montrer par deux méthodes différentes les égalités suivantes :

$$C_n^k = C_n^{n-k}, \quad C_n^k = C_{n-1}^{k-1} + C_{n-1}^k.$$

et

$$C_n^k = C_{n-2}^{k-2} + 2\ C_{n-2}^{k-1} + C_{n-2}^k.$$

Exercice 1.5. Soient $a, b \in \mathbb{R}, n \in \mathbb{N}$.

1. Montrer par récurrence la formule du binôme de Newton :

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

2. En déduire la valeur de

$$\sum_{k=0}^{n} (-1)^k C_n^k \quad \text{et} \quad \sum_{k=0}^{n} C_n^k$$

3. Montrer que si $card(\Omega) = n$ alors $card(\mathcal{P}(\Omega)) = 2^n$.
4. Montrer par deux méthodes différentes que pour tout $n \geq 1$:

$$\sum_{k=1}^{n} k \ C_n^k = n2^{n-1}.$$

(*On pourra utiliser la fonction* $f(x) = (1+x)^n$).

Exercice 1.6. Dans une classe de 30 étudiants :18 filles et 12 garçons. On veut former un comité composé de trois membres. Combien de comités peut-on former si :

a. aucune conditition n'est supposée.

b. le comité doit contenir les deux sexes.

c. le comité doit contenir au moins une fille.

Exercice 1.7. Soient $A, B \in \mathcal{P}(\Omega)$. Montrer les égalités suivantes :

- $\mathbb{1}_A = 1$ et $\mathbb{1}_\emptyset = 0$.
- $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B}$.
- $\mathbb{1}_{A\cup B} = \max(\mathbb{1}_A, \mathbb{1}_B)$.
- $\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$
- $\mathbb{1}_{A\cap B} = \mathbb{1}_A \mathbb{1}_B = \min(\mathbb{1}_A, \mathbb{1}_B)$

Chapitre 2

Espace de Probabilité

2.1 Espace probabilisable

2.1.1 Espace échantillon

Définition 2.1. On appelle expérience aléatoire ou épreuve, toute expérience dont le résultat est régi par le hasard l'orsqu'on répète l'expérience dans les mêmes conditions.

Exemple 2.1. Donnons quelques exemples simples d'expériences aléatoires :

1. le jet d'une pièce de monnaie et l'observation de la face supérieure.
2. Le jet d'un dé à six faces et l'observation de la face supérieure.
3. L'extraction d'une carte d'un jeu de 32 cartes.
4. La mesure de la durée de vie d'une batterie de téléphone portable.
5. La détermination du nombre de personnes arrivant à un guichet dans une période donnée.

Définition 2.2. On appelle l'ensemble des tous les résultats possibles d'une expérience aléatoire espace échantillon ou espace des épreuves. On le note Ω.

Remarque 2.1. Les espaces échantillons correspondent aux expériencs aléatoires citées dans l'exemple précédent sont respectivement : $\Omega_1 = \{\text{pile}, \text{face}\}$, $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$, $\Omega_3 =$ l'ensembles des 32 carte, $\Omega_4 = [0, +\infty[$, $\Omega_5 = \mathbb{N}$.

2.1.2 Algèbre-Tribu

Définition 2.3. Soit Ω un espace échantillon, une algèbre $\mathcal{A}$ sur Ω est une partie de $\mathcal{P}(\Omega)$ vérifiant :

i. $\Omega \in \mathcal{A}$

ii. $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$.

iii. $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Définition 2.4. Soit Ω un espace échantillon, une tribu ou une σ- algèbre $\mathcal{F}$ sur Ω est une partie de $\mathcal{P}(\Omega)$ vérifiant :

i. $\Omega \in \mathcal{A}$

ii. $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$.

iii. Si $\{A_k,\ k \in \mathbb{N}\}$ est une famille d'éléments de $\mathcal{F}$ alors

$$\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{F}.$$

Exemple 2.2. Soient Ω un espace échantillon, $A \in \mathcal{P}(\Omega)$, alors

- $\mathcal{A}_0 = \{\emptyset, \Omega\}$ est une algèbre sur Ω appelée l'algèbre triviale sur Ω.
- $\mathcal{P}(\Omega)$ est une algèbre sur Ω, elle est appelée l'algèbre grossière sur Ω.
- Soit $A \subset \Omega$, posons : $\sigma(A) = \{\emptyset, \Omega, A, \overline{A}\}$. Il est clair que $\sigma(A)$ est une algèbre sur Ω, elle est appelée l'algèbre engendrée pas A.
- Posons $\Omega = \{a, b, c\}$ et $\mathcal{A} = \{\emptyset, \Omega, \{a, b\}, \{a, c\}, \{c\}\}$, alors $\mathcal{A}$ n'est pas une algèbre sur Ω puisque $\overline{\{a, c\}} = \{b\} \notin \mathcal{A}$.

Remarque 2.2. Il n'est pas difficile de voir que

- Une tribu est forcément une algèbre mais la réciproque est fausse en général. (*voir exercice 2.9*)
- Une algèbre (et à fortiori une tribu) contient toujours l'ensemble vide.
- Une algèbre est stable par intersection finie. (*Voir exercice 2.7*)
- Une tribu est stable par intersection infinie dénombrable. (*Voir exercice 2.7*)
- La réunion de deux algèbres (resp. de deux tribus) n'est pas une algèbre (resp. une tribu) en général. (*Voir exercice 2.8*)

Cependant on a le résultat suivant :

Proposition 2.1. L'intersection de deux algèbres (resp. de deux tribus) est une algèbre (resp. une tribu).

Définition 2.5. Soient Ω un espace échantillon, $\mathcal{F}$ une tribu sur Ω. Le couple $(\Omega, \mathcal{F})$ est appelé espace probabilisable ou espace mesurable.

2.2 Espace de probabilité

Dans tout ce qui suit $(\Omega, \mathcal{F})$ désigne un espace probabilisable.

Définition 2.6. Soient $A, B, A_1, A_2, ...A_n$ des éléments de $\mathcal{F}$.

a. On dit que A et B sont incompatibles ou disjoints si $A \cap B = \emptyset$.

b. On dit que $A_1, A_2, ...A_n$ sont deux à deux incompatibles si pour tout $1 \leq i \neq j \leq n : \; A_i \cap A_j = \emptyset$.

c. On dit que $A_1, A_2, ...A_n$ forment une partition de Ω, s'ils sont deux à deux incompatibles et si
$$\bigcup_{k=1}^{n} A_k = \Omega.$$

Nous pouvons à présent donner la définition d'une probabilité :

Définition 2.7. On appelle probabilité, toute application $I\!P : \Omega \longrightarrow [0, 1]$ vérifiant les deux propriétés suivantes :

i. $I\!P(\Omega) = 1$.

ii. σ-ADDITIVITÉ : Pour toute famille $\{A_k, \; k \geq 1\}$ d'éléments de $\mathcal{F}$ deux à deux incompatibles, on a
$$I\!P\Big(\bigcup_{k\geq 1} A_k\Big) = \sum_{k=1}^{\infty} I\!P(A_k). \tag{2.1}$$

Remarque 2.3. Cette définition mérite quelques explications :

1. La première condition est assez naturelle puisque par définition, Ω est l'ensemble de tous les résultats possibles.

2. La condition de σ-additivité est aussi naturelle, en effet si on considère le jet d'un dé équilibré, la probabilité d'avoir un 2 ou un 4 est $2/6 = 1/6+1/6$ donc la somme des probabilités d'avoir un 2 et un 4.
3. De cette définition, on peut comprendre l'utilité de la notion de tribu, introduite dans la définition 2.4, ainsi la tribu peut être considérée comme le domaine de définition d'une probabilité $I\!P$. En effet dans cette dernière définition on peut parler de $I\!P(A_n)$ puisque $A_n \in \mathcal{F}$, mais on ne pouvait pas écrire $I\!P(\bigcup_{n\geq 1} A_n)$ si $\mathcal{F}$ n'était pas une tribu. (*à méditer*)
4. On peut se demander pourquoi ne pas définir une probabilité $I\!P$ sur $\mathcal{P}(\Omega)$ tout simplement ? La réponse est que lorsque Ω est un ensemble fini, la tribu $\mathcal{F}$ sera souvent $\mathcal{P}(\Omega)$, mais lorsque Ω est un ensemble infini, ceci n'est pas possible en général, ces considérations sont purement théoriques et sortent du cadre de ce polycopié.

Remarque 2.4. Il faut bien noter que pour la condition de σ-additivité exigée dans la définition précédente, on demande que les ensembles $\{A_k,\ k \geq 1\}$ soient deux à deux incompatibles, en effet si cette condition n'est pas vérifiée, l'équation (2.1), peut ne pas être satisfaite, pour le voir il suffit de considérer l'expérience aléatoire qui consiste à jeter un dé équilibré et à observer la face supérieure. Soient :

A : "avoir un chiffre pair"

B : "avoir un chiffre supérieur ou égale à 3".

On a $A = \{2,4,6\},\ B = \{3,4,5,6\}$, d'où $A \cup B = \{2,3,4,5,6\}$. Ainsi $I\!P(A) = 1/2,\ I\!P(B) = 2/3$ et $I\!P(A\cup B) = 5/6$, par conséquent

$$I\!P(A\cup B) = \frac{5}{6} \neq \frac{7}{6} = I\!P(A) + I\!P(B).$$

.

Définition 2.8. Si $(\Omega, \mathcal{F})$ est un espace probabilisable, $I\!P$ une probabilité définie sur $(\Omega, \mathcal{F})$, alors le triplet $(\Omega, \mathcal{F}, I\!P)$ est appelé <u>espace de probabilité</u>.

Proposition 2.2. Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilité, $A, B \in \mathcal{F}$, alors

1. $\forall E \in \mathcal{F},\quad 0 \leq I\!P(E) \leq 1.$

2. $I\!P(\emptyset) = 0$.
3. $I\!P(\overline{A}) = 1 - I\!P(A)$.
4. <u>Croissance</u> : $I\!P$ est une application croissante :

$$A \subset B \implies I\!P(A) \leq I\!P(B). \tag{2.2}$$

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

Exemple 2.3. Soient $(\Omega, \mathcal{F})$ un espace probabilisable, $A, B \in \mathcal{F}$. Peut-on définir une probabilité vérifiant :

$$I\!P(A) = \frac{1}{3}, \quad I\!P(B) = \frac{3}{4} \text{ et } I\!P(A \cap B) = \frac{1}{2}?$$

<u>Solution</u>:

Il suffit de remarquer que $I\!P(A \cap B) = \frac{1}{2} > \frac{1}{3} = I\!P(A)$ ce qui est impossible car $A \cap B \subset A$ et ceci implique d'après le quatrième point de la proposition 3.2 que $I\!P(A \cap B)$ doit être inférieur ou égale à $I\!P(A)$.

Exemple 2.4. Soient $(\Omega, \mathcal{F})$ un espace probabilisable, $A, B \in \mathcal{F}$. Peut-on définir une probabilité $I\!P$ vérifiant :

$$I\!P(A) = \frac{1}{2}, \quad I\!P(B) = \frac{7}{8} \text{ et } I\!P(A \cap B) = \frac{1}{4}.$$

<u>Solution</u>:

D'après la proposition 3.2, on a

$$I\!P(A \cup B) = \frac{1}{2} + \frac{7}{8} - \frac{1}{4} = \frac{9}{8} > 1,$$

ce qui est impossible d'après la définition d'une probabilité.

Remarque 2.5. Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilité, $A, B \in \mathcal{F}$, alors

- L'ensemble vide est appelé l'évènement impossible.
- Si $I\!P(A) = 0$, on dit que A est un évènement presque impossible.
- Ω est appelé l'évènement certain.
- Si $I\!P(B) = 1$, on dit que B est un évènement presque certain.

Dans la définition de la σ-additivité (définition 2.7), on exige que les évènements $\{A_k,\ k \in \mathbb{N}\}$ soient deux à deux incompatibles, pour pouvoir calculer la probabilité de la réunion de plusieurs évènements. Dans le cas où cette condition n'est pas vérifiée, on a seulement une inégalité, dite inégalité de Boole :

Proposition 2.3. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\{A_k,\ k \in \mathbb{N}\}$ une suite d'éléments de $\mathcal{F}$, alors

$$\mathbb{P}\Big(\bigcup_{k\geq 1} A_k\Big) \leq \sum_{k=1}^{\infty} \mathbb{P}(A_k). \tag{2.3}$$

Une probabilité possède aussi une propriété qui ressemble à la continuité d'une fonction et qui nous sera utile pour la suite :

Proposition 2.4. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité.

- Si $\{A_n,\ n \in \mathbb{N}\}$ est une suite croissante d'éléments de $\mathcal{F}$ (i. e $A_n \subset A_{n+1}$, pour tout $n \geq 1$), alors

$$\mathbb{P}\Big(\bigcup_{n\geq 1} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{2.4}$$

- Si $\{A_n,\ n \in \mathbb{N}\}$ est une suite décroissante d'éléments de $\mathcal{F}$ (i. e $A_n \supset A_{n+1}$, pour tout $n \geq 1$), alors

$$\mathbb{P}\Big(\bigcap_{n\geq 1} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n). \tag{2.5}$$

2.2.1 Le cas équiprobable

Supposons dans ce paragraphe que Ω est un ensemble fini avec $card(\Omega) = n$, on peut l'écrire : $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$. Supposons aussi que tous les évènements $\{\omega_k\}$ sont équiprobables ou uniformes i. e.

$$\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_2\}) = \cdots = \mathbb{P}(\{\omega_n\}). \tag{2.6}$$

En utilsant les propriétés de la probabilité $\mathbb{P}$ (définition 2.7), on a par la propiriété de σ-additivité :

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\bigcup_{k=1}^{n}\{\omega_k\}) = \sum_{k=1}^{n}\mathbb{P}(\{\omega_k\}) = n\mathbb{P}(\{\omega_1\}).$$

Par suite

$$\forall k \in \{1, 2, \dots, n\}, \qquad \mathbb{P}(\{\omega_k\}) = \frac{1}{n} \tag{2.7}$$

Soient à présent $1 \leq k \leq n,\ A \subset \Omega$, avec $card(A) = k$. Dans ce cas on peut écrire $A = \{\omega_1', \omega_2', \cdots, \omega_k'\}$, où chaque ω_i' est choisit dans l'ensemble Ω. Ainsi

$$\mathbb{P}(A) = \mathbb{P}(\bigcup_{j=1}^{k}\{\omega_j'\}) = \sum_{j=1}^{k}\mathbb{P}(\{\omega_j'\}) = k\mathbb{P}(\{\omega_1'\}) = \frac{k}{n}.$$

On a alors prouvé le résultat suivant :

Théorème 2.1. Dans le cas équiprobable, la probabilité d'un évènement A est donnée par

$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} \tag{2.8}$$

Afin d'illustrer ce résultat, donnons trois exemples :

Exemple 2.5. On jette deux dés équilibrés, et on observe les faces supérieures des deux dés. Notons par A l'évènement :

A : "la somme des deux chiffres est égale à 5".

Calculer $\mathbb{P}(A)$.

SOLUTION:

Soit Ω l'espace échantillon associé à cette expérience aléatoire, ainsi

$$\Omega = \{(x, y)/x, y \in \{1, 2, 3, 4, 5, 6\}\}.$$

Par suite $card(\Omega) = 6 \times 6 = 36$. De plus les deux dés sont équilibrés, donc chaque élément (x, y) de Ω a la même probabilité d'apparaître. On est par conséquent dans un cas équiprobable. D'autre part $A = \{(1, 4), (4, 1), (2, 3), (3, 2)\}$. Le théorème 2.6 donne :

$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{4}{36} = \frac{1}{9}.$$

Exemple 2.6. On forme un nombre de 4 chiffres choisis au hasard dans l'ensemble $\{1, 2, ..., 9\}$. Quelle est la probabilités que le nombre soit pair ?

<u>SOLUTION</u>:

Notons le nombre choisis par $abcd$ et soit Ω l'espace échantillon associé à cette expérience aléatoire, ainsi

$$\Omega = \{abcd/a, b, c, d \in \{1, 2, ...9\}\}.$$

D'où $card(\Omega) = 9 \times 9 \times 9 \times 9 = 9^4$. A présent, introduisons l'évènement B : "le nombre choisis est pair". Puisque le nombre est choisi au hasard, on est dans un cas équiprobable. D'autre part le nombre $abcd$ est pair si et seulement si $d \in \{2, 4, 6, 8\}$, donc $card(B) = 9 \times 9 \times 9 \times 4$. On a par le théorème

$$\mathbb{P}(B) = \frac{card(B)}{card(\Omega)} = \frac{9^3 \times 4}{9^4} = \frac{4}{9}.$$

Exemple 2.7. Une urne contient 10 boules : 3 noires, 4 blanches et 3 rouges. On tire de cette urne 3 boules. calculer la probabilité des évènements suivants :

- A : "Avoir exactemeent 3 boules blanches".
- B : "Avoir une boule de chaque couleur".
- C : "Avoir au moins une boule rouge".

<u>SOLUTION</u>:

Dans cet exemple on n'attache pas d'importance à l'ordre d'apparition des boules, on forme donc un sous ensemble composé de trois boules choisis dans un ensemnble de 10 boules. Par conséquent $card(\Omega) = C_{10}^3$.

- Pour le premier évènement, on veut 3 boules blanches, pour réaliser cet évènement, on doit choisir nos 3 boules parmi les 4 boules blanches, on a donc $card(A) = C_4^3$. Ainsi

$$\mathbb{P}(A) = \frac{card(A)}{card(\Omega)} = \frac{C_4^3}{C_{10}^3}.$$

– Pour l'évènement B, on veut avoir une boule de chaque couleur, donc on a 3 choix pour la noire, 4 pour la blanches et 3 pour la rouges, ainsi $card(B) = 3 \times 4 \times 3 = 36$. Donc

$$\mathbb{P}(B) = \frac{card(B)}{card(\Omega)} = \frac{36}{C_{10}^3}.$$

– Enfin pour le dernier évènement, même si on peut calculer $\mathbb{P}(C)$ directement, il est préférable de calculer d'abord $\mathbb{P}(\overline{C})$ qui est plus simple. En effet, on a $\overline{C}$: "Ne pas avoir de boule rouge", mais il y a 7 boules qui ne sont pas de couleur rouge, d'où $card(\overline{C}) = C_7^3$. Par conséquent

$$\mathbb{P}(C) = 1 - \mathbb{P}(\overline{C}) = 1 - \frac{card(\overline{C})}{card(\Omega)} = 1 - \frac{C_7^3}{C_{10}^3}.$$

2.2.2 Probabilité de la réunion

En calcul des probabilités, on est souvent amené à calculer la probabilité de l'union de plusieurs évènements, si ces évènements sont deux à deux incompatibles alors on peut utiliser directement la formule 2.1, dans le cas général, la formule suivante, dite FORMULE DE POINCARÉ, donne cette probabilité :

Théorème 2.2. Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité, $\{A_k,\ 1 \leq k \leq n\}$ une suite d'éléments de $\mathcal{F}$, alors on a

$$\mathbb{P}\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} (-1)^{k-1} S_k, \tag{2.9}$$

où

$$S_k = \sum_{i_1} \sum_{i_2} \cdots \sum_{i_k} \mathbb{P}\Big(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\Big),$$

où les sommes sont prises sur l'ensemble $\{1 \leq i_1 < i_2 < \ldots < i_k \leq n\}$.

Remarque 2.6. Dans cette formule le terme S_k représente la somme des probabilités de toutes les intersections possibles de k- évènements choisis parmis les évènements $A_1, A_2, \ldots, A_n$.

Appliquons cette formule pour $n = 2$ et $n = 3$:

– Pour $n = 2$, on a deux évènements A_1, A_2 et donc deux sommes à calculer S_1, S_2 :

$$S_1 = \mathbb{P}(A_1) + \mathbb{P}(A_2), \quad S_2 = \mathbb{P}(A_1 \cap A_2),$$

d'où

$$\mathbb{P}(A_1 \cup A_2) = \sum_{k=1}^{2} (-1)^{k-1} S_k = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2).$$

– Pour $n = 3$, on a trois évènements A_1, A_2, A_3 et donc trois sommes à calculer S_1, S_2, S_3 :

$$S_1 = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3), \quad S_2 = \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_3) + \mathbb{P}(A_2 \cap A_3),$$

$$S_3 = \mathbb{P}(A_1 \cap A_2 \cap A_3),$$

d'où

$$\begin{aligned}
\mathbb{P}(A_1 \cup A_2 \cup A_3) &= \sum_{k=1}^{3} (-1)^{k-1} S_k \\
&= S_1 - S_2 + S_3 \\
&= \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) \\
&\quad - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3).
\end{aligned}$$

Exemple 2.8. Quatre personnes entrent dans un restaurent pour diner muni chacun de son parapluie noir, en sortant chacun prend un parapluie au hasard. Calculer la probabilité qu'au moins une personne récupère son parapluie.

SOLUTION:

introduisons les évènements :

E : "Au moins une personne récupère son parapluie"

A_i : "La ième personne récupère son parapluie", $\quad i \in \{1, 2, 3, 4\}$

Dans cet exemple, Ω est l'ensemble de façons de distribuer les quatres parapluies sur leur propriétaires d'une manière aléatoire, ainsi $card(\Omega) = 4! = 24$. D'autre part $card(A_i)$ est le nombre de manières de distribuer les parpluies sur les quatres personnes en donnant à la ième personne le sien, ainsi on commence par donner à la ième personne son parapluie, ensuite on distribue les trois parapluie restant sur les trois autres personnes, par conséquent $card(A_i) = 3! = 6$. De même $card(A_i \cap A_j) = 2$ et $card(A_i \cap A_j \cap A_k) = card(A_1 \cap A_2 \cap A_3 \cap A_4) = 1$.

$$I\!P(A_i) = \frac{card(A_i)}{card(\Omega)}$$

Il est clair que $E = \cup_{i=1}^{4} A_i$, d'autre part, on a par le théorème 2.2 :

$$\begin{aligned} I\!P(E) &= I\!P(\cup_{i=1}^{4} A_i) \\ &= \sum_{k=1}^{4} (-1)^{k-1} S_k \\ &= S_1 - S_2 + S_3 - S_4 \end{aligned}$$

où

$$S_1 = \sum_{i=1}^{4} I\!P(A_i), \quad S_2 = \sum_{1 \leq i \neq j \leq 4} I\!P(A_i \cap A_j),$$

$$S_3 = I\!P(A_1 \cap A_2 \cap A_3) + I\!P(A_1 \cap A_3 \cap A_4) + I\!P(A_1 \cap A_2 \cap A_4) + I\!P(A_2 \cap A_3 \cap A_4)$$

et

$$S_4 = I\!P(A_1 \cap A_2 \cap A_3 \cup A_4).$$

2.3 Probabilité conditionnelle

Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilité, $B \in \mathcal{F}$ avec $I\!P(B) > 0$.

Définition 2.9. Pour $A \in \mathcal{F}$, on désigne par $I\!P(A/B)$ la probabilité de A sachant B, elle est définie par

$$I\!P(A/B) = \frac{I\!P(A \cap B)}{I\!P(B)}. \tag{2.10}$$

Remarque 2.7. Soient A, B deux évènements de Ω, $I\!P(A/B)$ se lit la probabilité de A sachant B ou la probabilité conditionnelle de A par rapport à B.

La probabilité condionnelle par rapport à un évènement permet d'introduire une nouvelle probabilité sur l'espace probabilisable $(\Omega, \mathcal{F})$, en effet :

Proposition 2.5. L'application $I\!P(\cdot/B)$ définie de $\mathcal{F}$ vers $[0,1]$ par

$$\forall A \in \mathcal{F}, \qquad I\!P(A/B) = \frac{I\!P(A \cap B)}{I\!P(B)}$$

est une probabilité sur $(\Omega, \mathcal{F})$.

Remarque 2.8. Soient $C, D \in \mathcal{F}$. Cette proposition signifie en particulier que

a. $I\!P(\Omega/B) = 1$.

b. $I\!P(C \cup D/B) = I\!P(C/B) + I\!P(D/B) - I\!P(C \cap D/B)$.

c. Si C et D sont incompatibles, alors $I\!P(C \cup D/B) = I\!P(C/B) + I\!P(D/B)$.

d. $I\!P(\overline{C}/B) = 1 - I\!P(C/B)$.

Exemple 2.9. On jette successivement deux dés équilibrés. calculer la probabilité des évènements suivants :

1. A : "La somme des chiffres sur les deux dés est un nombre pair".
2. B : "La somme des chiffres sur les deux dés est un nombre pair
   sachant que le premier dé a donné le chiffre 5".

SOLUTION:

Dans cet exemple $\Omega = \{(x,y)/x, y \in \{1,2,3,4,5,6\}\}$. D'autre part, la somme de deux chiffres est paire si et seulement si les deux chiffres sont pairs ou impairs, d'où

$$A = \{(x,y)/x, y \in \{1,3,5\}\} \cup \{(x,y)/x, y \in \{2,4,6\}\}.$$

Par conséquent, $card(A) = 9 + 9 = 18$ et on a

$$I\!P(A) = \frac{18}{36} = \frac{1}{2}.$$

Pour le second évènement introduisons l'évènement C : "le premier dé a donné le chiffre 5". On a ainsi $A \cap C = \{(5,y)/y \in \{1,3,5\}\}$ et on a $card(A \cap C) = 3$. Donc

$$I\!P(B) = I\!P(A/C) = \frac{I\!P(A \cap C)}{I\!P(C)} = \frac{3/36}{3/6} = \frac{1}{6}.$$

Remarque 2.9. Dans cet exemple, on a calculé dans les deux questions la probabilité que la somme des deux chiffres obtenus soit paire, mais à la difference de la première question où on n'avait aucune information, dans la deuxièmme question on savait que le premier dé a ramené le chiffre 5, ce qui explique la différence dans les résultats obtenus, et le fait que la probabilité calculée dans le second cas est plus petite que celle calculée au premier cas.

Exemple 2.10. On jette deux pièces de monnaie équilibrés. Calculer

1. La probabilité que les deux pièces ramènent pile, sachant que la première a ramené pile.
2. La probabilité que les deux pièces ramènent face, sachant qu'au moins l'une d'entre elle a ramené face.

<u>Solution</u>:

Pour $i \in \{1,2\}$, posons : A_i : "La ième pièce a ramené pile"

1. On cherche $I\!P(A_1 \cap A_2/A_1)$. On a

$$I\!P(A_1 \cap A_2/A_1) = \frac{I\!P(A_1 \cap A_2)}{I\!P(A_1)} = \frac{1/2 \times 1/2}{1/2} = \frac{1}{2}.$$

2. On cherche $I\!P(\overline{A_1} \cap \overline{A_2}/\overline{A_1} \cup \overline{A_2})$. On a

$$\begin{aligned} I\!P(\overline{A_1} \cap \overline{A_2}/\overline{A_1} \cup \overline{A_2}) &= \frac{I\!P\Big(\overline{A_1} \cap \overline{A_2} \cap (\overline{A_1} \cup \overline{A_2})\Big)}{I\!P(\overline{A_1} \cup \overline{A_2})} \\ &= \frac{I\!P(\overline{A_1} \cap \overline{A_2})}{1 - I\!P(\overline{A_1 \cap A_2})} \\ &= \frac{1/2 \times 1/2}{1 - (1/2 \times 1/2)} \\ &= \frac{1}{3}. \end{aligned}$$

Remarque 2.10. Soient à présent $A, C \in \mathcal{F}$ avec $I\!P(A) > 0$ et $I\!P(A \cap B) > 0$ D'après la définition de la probabilité conditionnelle, on a

$$I\!P(A/B) = \frac{I\!P(A \cap B)}{I\!P(B)},$$

d'où

$$\mathbb{P}(A \cap B) = \mathbb{P}(B/A)\mathbb{P}(A) = \mathbb{P}(A/B)\mathbb{P}(B). \tag{2.11}$$

De même

$$\mathbb{P}(C/A \cap B) = \frac{\mathbb{P}(C \cap A \cap B)}{\mathbb{P}(A \cap B)},$$

ainsi

$$\mathbb{P}(C \cap A \cap B) = \mathbb{P}(C/A \cap B)\mathbb{P}(A \cap B). \tag{2.12}$$

En combinant cette dernière équation avec (2.11), on obtient

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

En faisant un raisonnement par récurrence, on peut généraliser cette dernière formule à n évènements, ainsi

Proposition 2.6. Soient $n \geq 2$, $A_1, A_2, \ldots, A_n$ une famille d'éléments de $\mathcal{F}$, vérifiant $\mathbb{P}(\bigcap_{k=1}^{n-1} A_k) > 0$, alors

$$\mathbb{P}(\bigcap_{k=1}^{n} A_k) = \mathbb{P}(A_1)\mathbb{P}(A_2/A_1)\mathbb{P}(A_3/A_1 \cap A_2) \cdots \mathbb{P}(A_n/\bigcap_{k=1}^{n-1} A_k).$$

Remarque 2.11. Cette proposition est souvent utilisée lorsqu'on veut calculer la probabilité de l'intersection de plusieurs évènements obtenus en répétant une expérience aléatoire plusieurs fois, et lorsque l'espace échantillon change à chaque répétition.

Nous proposons un exemple simple pour illustrer cela :

Exemple 2.11. Une urne contient une boule blanche et une boule noire, on tire des boules de cette urne jusqu'à ce que la noire appraisse. A chaque fois qu'une boule blanche est tirée, on la remet dans l'urne avec une autre boule blanche. Calculer la probabilité que la boule noire n'apparaisse pas au cours des cinq premiers tirages.

<u>SOLUTION</u>:

Pour $i \in \{1, 2, 3, 4, 5\}$, posons :

B_i : "la boule tirée au ième tirage est blanche"

B : "la boule noire n'apparaisse pas au cours des cinq premiers tirages."

Il est facile de voir que

$$B = \bigcap_{i=1}^{5} B_i.$$

En appliquant la proposition précédente, on a n'apparaisse pas au cours des cinq premiers tirages.

$$\begin{aligned}
I\!P(B) &= I\!P\Big(\bigcap_{i=1}^{5} B_i\Big) \\
&= I\!P(B_1)I\!P(B_2/B_1)\cdots I\!P(B_5/B_1\cap B_2\cap B_3\cap B_4) \\
&= \frac{1}{2}\times\frac{2}{3}\times\frac{3}{4}\times\frac{4}{5}\times\frac{5}{6} \\
&= \frac{1}{6}.
\end{aligned}$$

2.3.1 Formule des probabilités totales

Dans certaines situations on est amené à calculer la probabilité d'un évènement A, qui peut se réaliser à travers plusieurs alternatives.

Soient $n \geq 2$, $A_1, A_2, \ldots, A_n$ une partition de Ω avec $I\!P(A_k) > 0$ pour tout $1 \leq k \leq n$. Si $A \in \mathcal{F}$, on peut écrire

$$I\!P(A) = I\!P(A\cap\Omega) = I\!P\Big(A\cap(\bigcup_{k=1}^{n} A_k)\Big) = I\!P\Big(\bigcup_{k=1}^{n}(A\cap A_k)\Big).$$

Mais les évènements $\{A\cap A_k,\ k = 1, 2, ..., n\}$ sont 2 à 2 incompatibles, ainsi

$$I\!P\Big(\bigcup_{k=1}^{n}(A\cap A_k)\Big) = \sum_{k=1}^{n} I\!P(A\cap A_k) = \sum_{k=1}^{n} I\!P(A_k)I\!P(A/A_k).$$

On obtient ainsi :

Théorème 2.3. Soient $A_1, A_2, \ldots, A_n$ un système complet de Ω, $A \in \mathcal{F}$, alors on a la FORMULE DE PROBABILITÉS TOTALES :

$$I\!P(A) = \sum_{k=1}^{n} I\!P(A_k)I\!P(A/A_k). \tag{2.13}$$

Exemple 2.12. Une zone sensible est couverte par un radar, le constructeur de ce radar affirme que si un avion est présent dans sa zone de couverture, ce radar le détecte avec une probabilité 0.99, alors qu'il peut signaler un objet sur son écran sans la présence d'avion avec une probabilité 0.1. La probabilité qu'un avion survol cette zone est 0.05.

a. Calculer la probabilité que le radar signale la présense d'un objet sur son écran.

b. Calculer la probabilité que le radar faille à sa mission.

SOLUTION:

Introdsuisons les évènements :

– A : "un avion est présent dans la zone de couverture".

– E : "le radar signale la présence d'un objet sur son écran".

a. On veut calculer $\mathbb{P}(E)$. On a par la formule de probabilité totale :

$$\begin{aligned} \mathbb{P}(E) &= \mathbb{P}(E/A)\mathbb{P}(A) + \mathbb{P}(E/\overline{A})\mathbb{P}(\overline{A}) \\ &= 0.99 \times 0.05 + 0.1 \times 0.95 \\ &= 0.1445 \end{aligned}$$

b. On cherche $\mathbb{P}(\overline{E} \cap A)$, il faut d'abord remarquer que par la proposition 2.5, l'application $B \mapsto \mathbb{P}(B/A)$ est une probabilité sur $(\Omega, \mathcal{F})$, ainsi

$$\mathbb{P}(\overline{E}/A) = 1 - \mathbb{P}(E/A)$$

D'autre part,

$$\begin{aligned} \mathbb{P}(\overline{E} \cap A) &= \mathbb{P}(\overline{E}/A)\mathbb{P}(A) \\ &= \Big[1 - \mathbb{P}(E/A)\Big]\mathbb{P}(A) \\ &= 0.01 \times 0.05 = 0.005. \end{aligned}$$

2.3.2 Formule de Bayes

Soient $A_1, A_2, \ldots, A_n$ une partition de Ω, $A \in \mathcal{F}$ et supposons que $\mathbb{P}(A) > 0$ et que $\mathbb{P}(A_k) > 0$ pour tout $1 \leq k \leq n$. Soit $1 \leq k \leq n$, on a par les équations

(2.11) :

$$I\!P(A \cap A_k) = I\!P(A/A_k)I\!P(A_k) = I\!P(A_k/A)I\!P(A).$$

Par conséquent

$$I\!P(A_k/A) = \frac{I\!P(A/A_k)I\!P(A_k)}{I\!P(A)}$$

En faisant appel à la formule des probabilités totales dans cette dernière égalité, on obtient :

Théorème 2.4. Soient $A_1, A_2, \ldots, A_n$ un système complet de Ω, $A \in \mathcal{F}$, avec $I\!P(A) > 0$, $I\!P(A_k) > 0$ pour tout $1 \leq k \leq n$. Alors

$$I\!P(A_k/A) = \frac{I\!P(A/A_k)I\!P(A_k)}{\sum_{k=1}^{n} I\!P(A/A_k)I\!P(A_k)} \qquad \text{(FORMULE DE BAYES)}.$$

Remarque 2.12. La formule de Bayes est appelée aussi la formule des causes, en effet la quantité $I\!P(A_k/A)$ donne la probabilité que l'évènement A s'est réalisé à travers A_k ou "à cause" de A_k.

Exemple 2.13. On effectue un test dans un grand élevage de bovins pour dépister une maladie. Ce test a permis de déceler 1.8% de cas atteints chez les mâles et 1.2% chez les femelles. Cet élevage contient 65% de femelles.

1. Quelle est la probabilité qu'un animal choisis au hasard dans cet élevage soit atteint de cette maladie.

2. L'animal choisi est atteint de cette maladie, quelle est la probabilité qu'il soit un mâle.

SOLUTION:

Introduisons les évènements suivants :

– A : "L'animal choisi est atteint de cette maladie".

– M : "L'animal choisi est un mâle".

Pour la première question, on cherche $I\!P(A)$. On a par la formule des probabilités totales :

$$\begin{aligned} I\!P(A) &= I\!P(A/M)I\!P(M) + I\!P(A/\overline{M})I\!P(\overline{M}) \\ &= (0.018)(0.35) + (0.012)(0.65) \\ &= 0.0141. \end{aligned}$$

Pour la deuxièmme question on cherche $I\!P(M/A)$. On a par la formule de Bayes :

$$I\!P(M/A) = \frac{I\!P(A/M)I\!P(M)}{I\!P(A)} = \frac{(0.018)(0.35)}{0.0141} = 0.4468.$$

2.4 Evènements indépendants

Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilité, $A, B \in \mathcal{F}$ avec $I\!P(B) > 0$.

Définition 2.10. On dit que les évènements A et B sont indépendants si

$$I\!P(A/B) = I\!P(A). \tag{2.14}$$

Remarque 2.13. Soient A et B deux évènements de Ω :

- Intuitivement, A et B sont indépendants si la réalisation de B n'influe pas sur la rélisation de A et vice versa.
- Si A et B sont indépendants et si $I\!P(A) > 0$, alors il n'est pas difficile de voir qu'on a aussi :

$$I\!P(B/A) = I\!P(B).$$

Une conséquence importante de la définition est :

Théorème 2.5. Deux évènements A et B sont indépendants si et seulement si

$$I\!P(A \cap B) = I\!P(A)I\!P(B) \tag{2.15}$$

Remarque 2.14. On peut prendre l'équation (2.15) comme définition pour l'indépendance de deux évènements, dans ce cas on est pas obligé de supposer que la probabilité de l'un des deux évènements est strictement positive.

Exemple 2.14. On jette une pièce de monnaie deux fois de suites et on considère les évènements :

- A : "On obtient pile au premier jet"
- B : "On obtient le même résultat dans les deux jets"
- C : "On obtient pile dans les deux jets"

Pour cette expérience on a $\Omega = \{(x,y)/x, y \in \{pile, face\}\}$, ainsi $card(\Omega) = 4$. D'autre part :

$$A = \{(pile, pile), (pile, face)\}, \quad B = \{(pile, pile), (face, face)\},$$

$$C = \{(pile, pile)\}, \quad A \cap B = \{(pile, pile)\}, \quad A \cap C = \{(pile, pile)\},$$

d'où

$$I\!P(A) = I\!P(B) = \frac{1}{2}, \quad I\!P(C) = I\!P(A \cap B) = I\!P(A \cap C) = \frac{1}{4}.$$

Par suite

$$I\!P(A \cap B) = I\!P(A) I\!P(B).$$

Ainsi A et B sont indépendants. Mais

$$I\!P(A \cap C) = \frac{1}{4} \neq \frac{1}{8} = I\!P(A) I\!P(C),$$

ce qui montre que A et C ne sont pas indépendants.

Proposition 2.7. Soient A, B deux évènements de Ω, les propriétés suivantes sont équivalentes :

- Les évènements A et B sont indépendants.
- Les évènements $\overline{A}$ et B sont indépendants.
- Les évènements A et $\overline{B}$ sont indépendants.
- Les évènements $\overline{A}$ et $\overline{B}$ sont indépendants.

En utilisant le théorème 2.5, on peut généraliser la notion d'indépendance à plusieurs évènements :

Définition 2.11. Soient $n \geq 2$, $A_1, A_2, \ldots, A_n$ des évènements d'un espace de probabilité $(\Omega, \mathcal{F}, I\!P)$. On dit que $A_1, A_2, \ldots, A_n$ sont indépendants si

$$I\!P\Big(\bigcap_{i=1}^{p} A_{k_i}\Big) = \prod_{i=1}^{p} I\!P(A_{k_i}), \tag{2.16}$$

pour tout $k_1, k_2, ..., k_p \in \{1, 2, ..., n\}$.

Remarque 2.15. D'après cette définition, pour montrer que n évènemente sont indépendants il faut vérifier que l'équation (2.16) est valide pour toutes les intersections possibles des ces évènements, ainsi trois évènements A, B et C sont indépendants si

$$\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B), \quad \mathbb{P}(A\cap C) = \mathbb{P}(A)\mathbb{P}(C), \quad \mathbb{P}(B\cap C) = \mathbb{P}(B)\mathbb{P}(C)$$

et

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

Exemple 2.15. Considérons l'expérience aléatoire qui consiste à jeter une pièce de monnaie deux fois de suite et d'observer les résultats obtenus. Soient A, B et C les évènements :

- A : " Le résultat du premier jet est pile "
- B : " Le résultat du deuxième jet est pile "
- C : " On obtient le même résultat dans les deux jets "

Ici $\Omega = \{(x, y),\ \text{avec}\ x, y \in \{pile, face\}\}$, donc $card(\Omega) = 4$. D'autre part

$$\mathbb{P}(C) = \text{ La probabilité d'avoir deux fois pile ou deux fois face } = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

d'où

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$$

d'un autre coté

$\mathbb{P}(A\ \cap\ B) =$ La probabilité d'avoir pile dans les deux jets
$\mathbb{P}(A\ \cap\ C) =$ La probabilité d'avoir pile dans les deux jets
$\mathbb{P}(B\ \cap\ C) =$ La probabilité d'avoir pile dans les deux jets.

Par suite

$$\mathbb{P}(A\ \cap\ B) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B), \quad \mathbb{P}(A\ \cap\ C) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(C)$$

et

$$\mathbb{P}(B\ \cap\ C) = \frac{1}{4} = \mathbb{P}(B)\mathbb{P}(C).$$

Ainsi

$$\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

Par conséquent, les évènements A, B et C ne sont pas indépendants.

2.5 Exercices

Exercice 2.1. Soit $(\Omega, \mathcal{F})$ un espace probabilisable, A, B et C des évènements de Ω. En utilisant les opérations $\cap, \cup$ et le passage au complémentaire, exprimer les évènements suivants :

- E_1 : Aucun des évènements A, B et C ne se réalise.
- E_2 : B se réalise mais pas A et C.
- E_3 : Exactement deux évènements se réalisent.
- E_4 : Au moins l'un des évènements se réalisent.
- E_5 : Au plus l'un des évènements se réalisent.
- E_6 : Exactement un évènement se réalise.

Exercice 2.2. Dans une tombola il y a 120 tickets pour gagner 3 lots différents, pour chaque lot il y a 4 cadeaux et donc 4 tickets gagnants. Une personne achète 3 tickets.

1. Quelle est la probabilité que cette personne gagne un seul cadeau ?
2. Quelle est la probabilité que cette personne gagne au moins un cadeau ?
3. Quelle est la probabilité que cette personne un cadeau de tois lots différents ?

Exercice 2.3. On jette 3 dés équilibrés et on note les résultats de ces trois jets : a, b et c.

i. Déterminer Ω et $card(\Omega)$.

ii. On forme alors l'équation (E) : $ax^2 + bx + c = 0$. Calculer les probabilités des évènements suivants :
   - A : "Les racines de l'équation (E) sont réelles".
   - B : "Les racines de l'équation (E) sont complexes".

Exercice 2.4.

Exercice 2.5.

Exercice 2.6. Montrer par récurrence la proposition 2.6.

Exercice 2.7. Soient Ω un espace échantillon, $\mathcal{A}$ une algèbre sur Ω et $\mathcal{F}$ une tribu sur Ω.

1. Montrer que si $n \in \mathbb{N}$ et $A_1, A_2, ..., A_n$ des éléments de $\mathcal{A}$, alors

$$\bigcup_{k=1}^{n} A_k \in \mathcal{A} \quad \text{et} \quad \bigcap_{k=1}^{n} A_k \in \mathcal{A}.$$

2. Montrer que si $\{A_1, A_2, ...,\}$ est une suite d'éléments de $\mathcal{F}$, alors

$$\bigcap_{k \geq 1} A_k \in \mathcal{F}.$$

Exercice 2.8. En utilisant les algèbres engendrées par deux ensembles différents, montrer que l'union de deux algèbres n'est pas nécessairement une algèbre.

Exercice 2.9. Posons

$$\mathcal{A} = \{A \in \mathcal{P}(\mathbb{R}),\ card(A) \text{ est fini où } card(\overline{A}) \text{ est fini}\}.$$

Montrer que $\mathcal{A}$ est une algèbre sur $\mathbb{R}$, mais pas une tribu.

Exercice 2.10. Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilités, $A, B \in \mathcal{F}$ avec $I\!P(A) = 1/3$, $I\!P(B) = 1/4$ et $I\!P(A \cap B) = 1/6$. Calculer

$$I\!P(\overline{A}),\ I\!P(\overline{A} \cup B),\ I\!P(A \cup \overline{B}),\ I\!P(A \cap \overline{B}),\ I\!P(\overline{A} \cup \overline{B}).$$

Exercice 2.11. Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilités, $A, B \in \mathcal{F}$ avec $I\!P(A) = \frac{3}{4}$ et $I\!P(B) = \frac{1}{3}$. Montrer que $\frac{1}{12} \leq I\!P(A \cap B) \leq \frac{1}{3}$.

Exercice 2.12. Soient $(\Omega, \mathcal{F}, I\!P)$ un espace de probabilités, $A, B, C \in \mathcal{F}$ avec $I\!P(A \cap B \cap C) > 0$. Montrer que

1. $I\!P(\overline{A} \cup \overline{B}) = 1 - I\!P(B) I\!P(A/B)$.
2. $\frac{I\!P(A/(A \cup B))}{I\!P(B/(A \cup B))} = \frac{I\!P(A)}{I\!P(B)}$.
3. $I\!P(A \cap B/(B \cup C)) = I\!P(A \cap B/B) I\!P(B/(B \cup C))$.
4. $I\!P(A/B) I\!P(B/C) I\!P(C/A) = I\!P(B/A) I\!P(C/B) I\!P(A/C)$.
5. $I\!P(B \cap C/A) = I\!P(C/A) I\!P(B/(A \cap C))$.

Exercice 2.13. Dans une université 70% sont des garçons. 60% des garçons fument ainsi que 40% des filles. les étudiants sont enregistrés par un numéro d'inscription. Un numéro est tiré au hasard.

1. Quelle est la probabilité que le numéro tiré corresponde à une personne qui fume ?
2. Si on sait que le numéro tiré correspond à une personne qui fume, quelle est la probabilité que ce soit une fille ?

Exercice 2.14. La ville de Strasbourg contient 82% d'alsaciens et 18% de personnes d'origines non alsaciènnes. 25% des alsaciens parle allemand, 5% de personnes d'origine non alsaciènne parle allemand. Un touriste allemand est en visite à Strasbourg, il demande à un strasbourgeois choisis au hasard de lui indiquer le chemin du parlement européen.

1. Quelle la probabilité que ce Strasbourgeois parle allemand ?
2. La personne choisis parle effectivement allemand, quelle est la probabilité quelle soit alsaciènne ?

Exercice 2.15. Une école prestigieuse exige de ces candidats de passer un test avant d'accepter leurs candidature. Un bon candidat a 85% de chance de réussir le test alors qu'un candidat faible n'a que 15% de chance de réussir le test. Des études statistiques ont montré qu'il y a en moyenne 40% de bon candidats.

1. Quelle est la probabilité qu'un candidat chosis au hasard, réussi le test.
2. Un candidat a échoué au test, quelle est la probabilité qu'il soit bon.
3. Quelle est la proportion des bons candidats qui ont réussi au test.

Exercice 2.16. Soient $(\Omega, \mathcal{F}, I\!\!P)$ un espace de probabilités, $A, B, C \in \mathcal{F}$ avec $I\!\!P(C) > 0$ et $A_1, A_2, \cdots, A_n$ des éléments de $\mathcal{F}$.

1. Montrer que $I\!\!P(A \cup B/C) = I\!\!P(A/C) + I\!\!P(B/C) - I\!\!P(A \cap B/C)$.
2. Montrer que si $A_1, A_2, \cdots, A_n$ sont deux à deux incompatibles, alors

$$I\!\!P(\bigcup_{k=1}^{n} A_k \; /C) = \sum_{k=1}^{n} I\!\!P(A_k/C).$$

3. En déduire que l'application $I\!\!P(\cdot/B)$ définie de $\mathcal{F}$ vers $[0,1]$ par

$$I\!\!P(A/B) = \frac{I\!\!P(A \cap B)}{I\!\!P(B)}$$

est une probabilité sur $(\Omega, \mathcal{F})$.

Exercice 2.17. **Soient $0 < p < 1$, $A_1, A_2, ...A_n$ des évènements indépendants avec $P(A_1) = p$. Calculer en fonction de p la probabilité des évènements suivants :

- A : "Aucun évènement ne se réalise".
- B : "Au moins un évènement se réalise".

Exercice 2.18. Soient A, B deux évènements de Ω. Montrer que les propriétés suivantes sont équivalentes :

- Les évènements A et B sont indépendants.
- Les évènements $\overline{A}$ et B sont indépendants.
- Les évènements A et $\overline{B}$ sont indépendants.
- Les évènements $\overline{A}$ et $\overline{B}$ sont indépendants.

Exercice 2.19. Quatre personnes prennent un ascenseur dans un immeuble de quatre étages, chaque personne choisi de descendre à un étage indépendamment des autres. Calculer les probabilités suivantes :

- Les quatres personnes descendent au même étage.
- Les quatres personne descendent à des étages différents.
- Trois personnes descendent au même étage et la quatrième à un autre étage.

Exercice 2.20. Soient $(\Omega, \mathcal{F}, I\!\!P)$ un espace de probabilités, A, B deux éléments de $\mathcal{F}$. Montrer que

$$|I\!\!P(A) - I\!\!P(B)| \leq I\!\!P(A \Delta B)$$

Soient $A_1, \cdots, A_n, B_1, \cdots, B_n$ des éléments de $\mathcal{F}$ vérifiant $A_k \subset B_k$ pour $1 \leq k \leq n$. Montrer que

1. $$I\!\!P(\bigcup_{k=1}^{n} B_k) - I\!\!P(\bigcup_{k=1}^{n} A_k) \leq \sum_{k=1}^{n} (I\!\!P(B_k) - I\!\!P(A_k))$$

2.
$$\mathbb{P}(\bigcap_{k=1}^{n} A_k) \geq 1 - \sum_{k=1}^{n} \mathbb{P}(\overline{A_k})$$

Exercice 2.21. On fixe $n \geq 1$ et on choisit dans $\{1, 2, \cdots n\}$ un nombre d'une manière équiprobable. Si $p \leq n$ on note

$$A_p : \text{ "le nombre choisi est divisible par } p\text{".}$$

1. Calculer $\mathbb{P}(A_p)$ lorsque p est un diviseur de n.
2. Si $p_1, p_2 \cdots p_k$ sont des diviseurs premiers de n, deux à deux distinct, montrer que les événements associés $(A_i)_{1\leq i \leq k}$ sont indépendants.
3. On appelle indicateur d'Euler, que l'on note $\varphi(n)$, le nombre d'entiers strictement inférieurs à n et premiers avec n. Montrer la formule d'Euler :
$$\frac{\varphi(n)}{n} = \prod_{p \ premier, \ p \backslash n} (1 - \frac{1}{p})$$

Chapitre 3

Variables aléatoires réelles

Tout au long de ce chapitre $(\Omega, \mathcal{F}, I\!P)$ désignera un espace de probabilité.

3.1 Définition-Exemples

Nous commençons ce paragraphe par définir une tribu très importante en théorie de probabilité, la tribu boréliènne. pour ce faire introduisons d'abord la notion de tribu engendrée :

Définition 3.1. Soient $(\Omega, \mathcal{F})$ un espace probabilisable, $\mathcal{C} \subset \mathcal{P}(\Omega)$. On appelle tribu engendrée par $\mathcal{C}$, la plus petite tribu contenant $\mathcal{C}$.

Remarque 3.1. La tribu engendrée par une partie $\mathcal{C}$ de $\mathcal{P}(\Omega)$ existe toujours puisque $\mathcal{P}(\Omega)$ est une tribu contenant $\mathcal{C}$. Elle est construite en prenant l'intersection de toutes les tribus contenant $\mathcal{C}$.

Définition 3.2. La tribu boreliènne sur $\mathbb{R}$ (ou la tribu de Borel), notée $\mathcal{B}_{\mathbb{R}}$ est la tribu engendrée par la famille $\mathcal{C}_0 = \{]-\infty, x], x \in \mathbb{R}\}$.

Remarque 3.2. Les éléments de $\mathcal{B}_{\mathbb{R}}$ sont appelés les boréliens de $\mathbb{R}$.

Remarque 3.3. La tribu boreliènne sur $\mathbb{R}$ est construite à partir de la famille $\mathcal{C}_0$, en utilisant les opérations $\cup$, $\cap$ et le passage au complémentaire, ainsi $\mathcal{B}_{\mathbb{R}}$ contient les intervalles ouverts, fermés, bornés, non bornés, les points,..., etc. En effet, on peut voir par exemple que pour $x, y \in \mathbb{R}$ avec $x < y$, on a

– $[x, +\infty[= \overline{]-\infty, x]} \in \mathcal{B}_{\mathbb{R}}$.
– $[x, y[=]-\infty, y] \cap [x, +\infty[\in \mathcal{B}_{\mathbb{R}}$.

Définition 3.3. Soit $X : \Omega \longrightarrow \mathbb{R}$ une application. On dit que X est une variable aléatoire réelle (on note v. a. r) si

$$\forall A \in \mathcal{B}_{\mathbb{R}}, \qquad X^{-1}(A) \in \mathcal{F}. \tag{3.1}$$

Remarque 3.4. Il faut remarquer, que dans le cas général $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$, par conséquent, si X est une application de Ω vers $\mathbb{R}$ et si $A \in \mathcal{B}_{\mathbb{R}}$, on peut juste affirmer que $X^{-1}(A) \in \mathcal{P}(\Omega)$. Pour que X soit une variable aléatoire on demande que $X^{-1}(A)$ soit dans $\mathcal{F}$. (*à méditer*)

Exemple 3.1. Donnons quelques exemples simples :

1. L'application constante : Soient $\lambda \in \mathbb{R}$, X l'application constante définie sur Ω par $X(\omega) = \lambda$, pour tout $\omega \in \Omega$, alors X est une v. a. r.
   En effet soit $A \in \mathcal{B}_{\mathbb{R}}$, on a

$$X^{-1}(A) = \begin{cases} \Omega & \text{si } \lambda \in A \\ \emptyset & \text{si } \lambda \in \overline{A} \end{cases}$$

   d'où $X^{-1}(A) \in \mathcal{F}$ puisqu'une tribu contient toujours Ω et l'ensemble vide.
2. La fonction indicatrice : Soient $A \in \mathcal{F}$, $\mathbb{1}_A$ la fonction indicatrice de A (*voir ? ?*). Posons $Y = \mathbb{1}_A$, alors Y est une v. a. r. En effet soit $A \in \mathcal{B}_{\mathbb{R}}$, on a

$$Y^{-1}(A) = \begin{cases} \Omega & \text{si} \quad \{0,1\} \subset A \\ \emptyset & \text{si} \quad \{0,1\} \subset \overline{A} \\ A & \text{si } 1 \in A, \ 0 \in \overline{A} \\ \overline{A} & \text{si } 0 \in A, \ 1 \in \overline{A} \end{cases}$$

   d'où $Y^{-1}(A) \in \mathcal{F}$.
3. Supposons à présent que $\Omega = \{a, b, c, d\}$, $\mathcal{F} = \{\emptyset, \Omega, \{a\}, \{b, c, d\}\}$ et soit Z l'application définie sur Ω par $Z(a) = 1$, $Z(b) = Z(c) = 0$ et $Z(d) = 50$. On a alors

$$Z^{-1}(0) = \{b, c\} \notin \mathcal{F},$$

ainsi Z n'est pas une v. a. r.

Sur ces trois exemples simples, on a pu déterminer à partir de la définition si les applications proposées était des v. a. r ou pas, en général il est difficile d'utiliser cette définition pour savoir si une application donnée est une v. a. r ou pas. La difficulté réside dans le fait qu'on doit vérifier la condition (3.1) pour tous les boréliens de $\mathbb{R}$! Or un borélien de $\mathbb{R}$ peut être un ensemble assez compliqué, c'est pourquoi on a besoin d'un critère plus pratique pour déterminer si une application donnée est une v. a. r ou pas. Le résultat suivant donne ce critère.

Théorème 3.1. Une application $X : \ \Omega \longrightarrow \mathbb{R}$ est une variable aléatoire réelle si et seulement si

$$\forall x \in \mathbb{R}, \qquad X^{-1}(]-\infty, x]) \in \mathcal{F} \tag{3.2}$$

Remarque 3.5. Sous la lumière de ce résultat, on peut affirmer que

1. Pour prouver qu'une application est une v. a. r, il suffit de vérifier la condition (3.1) pour les boréliens de la forme $]-\infty, x]$, en lieu et place d'un borélien quelconque.
2. Il est évident que La condition (3.1) implique (3.2), l'implication réciproque est dûe au fait que la tribu boréliènne sur $\mathbb{R}$ est générée par les intervalles de type $]-\infty, x]$ et à un résultat classique de la théorie de probabilité appelé Théorème de la classe monotone.

`Notation` : Soient X une variable aléatoire réelle, $A \in \mathcal{B}_{\mathbb{R}}$. A partir de maintenant, la notation $I\!P(\{X \in A\})$ désignera $I\!P(\{\omega \in \Omega, \ X(\omega) \in A\})$.

Définition 3.4. Soit X une v. a. r définie sur Ω, la <u>fonction de répartition</u> F_X associée à X est définie par

$$\begin{aligned} F_X : \quad \mathbb{R} &\longrightarrow [0,1] \\ x &\longmapsto F_X(x) \end{aligned}$$

avec

$$F_X(x) = I\!P\Big(X^{-1}(]-\infty, x])\Big) = I\!P\Big(X \le x\Big).$$

Exemple 3.2. Déterminons les fonctions de répartition associées aux deux v. a. r de l'exemple 3.1.

1. Soit $\lambda \in \mathbb{R}$ et posons $X(\omega) = \lambda$ pour tout $\omega \in \Omega$. On alors

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

2. Soient $A \in \mathcal{F}$ et posons $Y = \mathbb{1}_A$, la fonction indicatrice de A. On a

$$F_Y(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - \mathbb{P}(A) & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

La fonction de répartition associée à une variable aléatoire réelle jouit des propriétés suivantes :

Proposition 3.1. Soient X une v. a. r, F_X sa fonction de répartition, alors

1. <u>L'IMAGE DE F_X</u> :
$$\forall x \in \mathbb{R}, \quad 0 \leq F_X(x) \leq 1.$$
2. <u>LE COMPORTEMENT À L'INFINI</u> :
$$\lim_{x \to -\infty} F_X(x) = 0 \quad \text{et} \quad \lim_{x \to +\infty} F_X(x) = 1.$$
3. <u>LA CROISSANCE</u> : La fonction F_X est croissante.
4. <u>LA CONTINUITÉ À DROITE</u> : La fonction F_X est continue à droite, i. e.
$$\forall a \in \mathbb{R}, \qquad \lim_{x \to a^+} F_X(x) = F_X(a).$$

La fonction de répartition est très utile pour calculer la probabilité attachée à un intervalle, en effet

Théorème 3.2. Soient $a, b \in \mathbb{R}$ avec $a < b$, X une v. a. r et F_X sa fonction de répartition, alors

1. $\mathbb{P}\Big(a < X \leq b\Big) = F_X(b) - F_X(a)$.
2. $\mathbb{P}\Big(a \leq X \leq b\Big) = F_X(b) - F_X(a) + \mathbb{P}(X = a)$.

3. $\mathbb{P}\Big(a < X < b\Big) = F_X(b) - F_X(a) - \mathbb{P}(X = b)$.
4. $\mathbb{P}\Big(a \leq X < b\Big) = F_X(b) - F_X(a) + \mathbb{P}(X = a) - \mathbb{P}(X = b)$.

PREUVE

Il est clair qu'il suffit de montrer la première égalité. Soient $a, b \in \mathbb{R}$ avec $a < b$, alors on a

$$X^{-1}\Big(]-\infty, b]\Big) = X^{-1}\Big(]-\infty, a]\Big) \ \cup \ X^{-1}\Big(]a, b]\Big),$$

de plus, les ensembles $X^{-1}\Big(]a, b]\Big)$ et $X^{-1}\Big(]-\infty, a]\Big)$ sont incompatibles, donc

$$\begin{aligned}
F_X(b) &= \mathbb{P}\Big(X^{-1}\Big(]-\infty, b]\Big)\Big) \\
&= \mathbb{P}\Big(X^{-1}\Big(]-\infty, a]\Big) \ \cup \ X^{-1}\Big(]a, b]\Big)\Big) \\
&= \mathbb{P}\Big(X^{-1}\Big(]-\infty, a]\Big)\Big) + \mathbb{P}\Big(X^{-1}\Big(]a, b]\Big)\Big) \\
&= \mathbb{P}\Big(X \leq a\Big) + \mathbb{P}\Big(a < X \leq b\Big) \\
&= F_X(a) + \mathbb{P}\Big(a < X \leq b\Big).
\end{aligned}$$

Par conséquent

$$\mathbb{P}\Big(a < X \leq b\Big) = F_X(b) - F_X(a).$$

Comme on l'a déja expliqué au début de ce chapitre, une v. a. r X permet de travailler sur $\mathbb{R}$, qui est un space plus "sympathique" que l'espace échantillon Ω, elle permet aussi de génerer une probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ appelée la loi de X, en effet :

Théorème 3.3. Soit X une v. a. r définie sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ et soit $\mathbb{P}_X$, l'application définie par

$$\begin{aligned}
\mathbb{P}_X : \quad \mathcal{B}_{\mathbb{R}} &\longrightarrow [0, 1] \\
A &\longmapsto \mathbb{P}_X(A)
\end{aligned}$$

avec

$$\forall A \in \mathcal{B}_{\mathbb{R}}, \qquad I\!\!P_X\big(A\big) = I\!\!P\Big(X^{-1}(A)\Big) = I\!\!P\Big(\omega \in \Omega,\ X(\omega) \in A\Big).$$

$I\!\!P_X$ est une probabilité sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, elle est appelée la loi de X.

Remarque 3.6. Il faut bien noter la différence entre les probabilités $I\!\!P$ et $I\!\!P_X$, la première est définie sur $(\Omega, \mathcal{F})$, c'est la probabilité de référence et elle est indépendante de la v. a. r X, alors que la seconde est construite sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ en faisant appel à la variable X.

Les variables aléatoires réelles se décomposent en deux grandes familles : les variables aléatoires discrètes et les variables aléatoires absolument continues, nous allons à présent étudier ces deux familles.

3.2 Variables aléatoires discrètes

Définition 3.5. On dit qu'un sous ensemble de $\mathbb{R}$ est discret si il est fini ou infini dénombrable.

Définition 3.6. Soit X une v. a. r définie sur un espace probabilisable $(\Omega, \mathcal{F})$. On dit que X est une v. a. r discrète si $X(\Omega)$ est un ensemble discret.

Le sous ensemble $X(\Omega)$ est appelé le support de X, c'est l'ensemble qui contient les valeurs prises par la v. a. r X avec une probabilité strictement positive.

Remarque 3.7. Une variable aléatoire discrète est tout simplement une application qui prend des valeurs discrètes $\{x_1, x_2, \dots,\ x_n, \dots\}$, ainsi on arrive à

Définition 3.7. Soit X une v. a. r discrète, on appelle fonction de masse de X, la fonction définie par

$$\begin{aligned} p_X :\ \ & \mathbb{R} \longrightarrow [0,1] \\ & x \longmapsto p_X(x) \end{aligned}$$

avec

$$p_X(x) = \begin{cases} I\!\!P(X = x) & \text{si } x \in X(\Omega) \\ 0 & \text{sinon} \end{cases}$$

Exemple 3.3. Déterminons les fonctions de masse associées aux deux variables aléatoires discrètes de l'exemple 3.1.

1. Soient $\lambda \in \mathbb{R},\ X(\omega) = \lambda$ pour tout $\omega \in \Omega$, alors X est une v. a. r discrète, $X(\Omega) = \{\lambda\}$ et sa fonction de masse est donnée par

$$p_X(x) = \begin{cases} 1 & \text{si } x = \lambda \\ 0 & \text{sinon} \end{cases}$$

2. Soient $A \in \mathcal{F},\ Y = \mathbb{1}_A$, la fonction indicatrice de A, alors la fonction de masse de Y est donnée par

$$p_X(x) = \begin{cases} \mathbb{P}(A) & \text{si } x = 1 \\ 1 - \mathbb{P}(A) & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases}$$

On a vu comment calculer la fonction de masse pour une variable aléatoire donnée, inversement on peut se demander sous quelles conditions, une fonction $p(x)$ est la fonction de masse d'une certaine variable aléatoire ? Le résultat suivant donne une réponse à cette question :

Proposition 3.2. Soient $\mathfrak{N}$ un sous ensemble discret de $\mathbb{R},\ p$ une fonction réelle définie sur $\mathfrak{N}$, nulle ailleurs et vérifiant :

i. Pour tout $x \in \mathbb{R},\ 0 \leq p(x) \leq 1$.

ii. $\sum_{k\in\mathfrak{N}} p(k) = 1$

Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, une variable aléatoire X définie sur cet espace telle que p soit la fonction de masse de X, i. e. $p_X(x) = p(x)$, pour tout $x \in \mathbb{R}$.

Exemple 3.4. La fonction p définie par

$$p(x) = \begin{cases} \frac{6}{\pi^2 k^2} & \text{si } x = k \in \mathbb{N}^* \\ 0 & \text{sinon} \end{cases}$$

est la fonction de masse d'une variable aléatoire car

$$\forall x \geq 1, \quad 0 \leq p(x) \leq 1 \qquad \text{et} \qquad \sum_{k\geq 1} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Exemple 3.5. Soient $n \geq 1, C > 0$ et p la fonction réelle définie par

$$p(x) = \begin{cases} C\frac{x}{n(n+1)} & \text{si } x \in \{1, 2, \dots, n\} \\ 0 & \text{sinon} \end{cases}$$

Sous quelles conditions, p est-elle une fonction de masse ?

<u>SOLUTION</u>:

La première chose à vérifier est la condition $0 \leq p(x) \leq 1$. Un calcul simple montre que cette condition est satisfaite si $C \leq n+1$. D'autre part on doit avoir aussi $\sum_{x\in\mathbb{N}} p(x) = 1$. Par conséquent

$$1 = \sum_{x\in\mathbb{N}} p(x) = \frac{C}{n(n+1)} \sum_{x=1}^{n} x = \frac{C}{n(n+1)} \times \frac{n(n+1)}{2} = \frac{C}{2}.$$

Donc, ci $C = 2$ les deux conditions sont vérifiées et p est bien une fonction de masse.

A présent déterminons la relation entre la fonction de masse et la fonction de répartition dans le cas d'une variable aléatoire discrète.

Théorème 3.4. Soient X une v. a. r discrète, F_X sa fonction de répartition et p_X sa fonction de masse. Pour $x \in X(\Omega)$, posons $F_X(x^-) = \lim_{t\to x^-} F_X(t)$. Alors

- Pour tout $x \in \mathbb{R}$,
$$p_X(x) = F_X(x) - F_X(x^-).$$
- Pour tout $x \in \mathbb{R}$,
$$F_X(x) = \sum_{k\leq x} p_X(k).$$

3.2.1 Espérance

Définition 3.8. On appelle espérance mathématique ou moyenne d'une variable aléatoire réelle discrète X, la quantité notée $\mathbb{E}(X)$ et définie par

$$\mathbb{E}(X) = \sum_{k \in X(\Omega)} k p_X(k).$$

où p_X est la fonction de masse de X.

Remarque 3.8. Soit X une v. a. r discrète.

i. L'espérance de X est définie par une somme qui peut être infinie, c'est pour cette raison qu'elle appartient à $\mathbb{R} \cup \{-\infty, +\infty\}$ en général.

ii. Lorsque $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est centrée.

iii. Si $X(\Omega)$ est un ensemble fini, alors il n'est pas difficile de montrer que

$$\min_{\omega \in \Omega} X(\omega) \leq \mathbb{E}(X) \leq \max_{\omega \in \Omega} X(\omega).$$

Proposition 3.3. Soient X une variable discrète, p_X sa fonction de masse et $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue. Posons $Y = g(X)$, alors Y est encore une v. a. r discrète et

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \sum_{k \in X(\Omega)} g(k) p_X(k).$$

Définition 3.9. Soit X une variable aléatoire réelle discrète. On dit que X est intégrable si $\mathbb{E}(|X|) < \infty$.

L'espérance d'une v. a. r, jouit d'une propriété très utile dans les calculs, elle est linéaire :

Proposition 3.4. Soient X, Y deux variables aléatoire discrète, $\alpha,\ \beta \in \mathbb{R}$, alors

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

ii. Si $\mathbb{E}(|X|) = 0$ alors $\mathbb{P}(X = 0) = 1$.

3.2.2 Variance et Ecart-type

Définition 3.10. Soit X une variable aléatoire réelle discrète.

– On appelle variance de X, la quantité notée $Var(X)$ et définie par

$$Var(X) = \mathbb{E}\Big[\Big(X - \mathbb{E}(X)\Big)^2\Big] = \sum_{k \in X(\Omega)} \Big(k - \mathbb{E}(X)\Big)^2 p_X(k).$$

– On appelle écart-type de X, la quantité notée $\sigma(X)$ et définie par

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

Remarque 3.9. La variance d'une v. a. r discrète est définie par la somme (éventuellement infinie) de quantités positives, par conséquent $Var(X) \in \mathbb{R}^+ \cup \{+\infty\}$ en général.

En développant le carré dans la définition de la variance, on obtient une nouvelle formule plus utile dans les calculs, en effet :

Proposition 3.5. Si X est une variable aléatoire discrète, alors

$$Var(X) = \mathbb{E}(X^2) - \Big(\mathbb{E}(X)\Big)^2 = \sum_{k \in X(\Omega)} k^2 p_X(k) \;-\; \Big(\mathbb{E}(X)\Big)^2.$$

Contrairement à l'espérance mathématique, la variance n'est pas linéaire, elle est cependant invariante par translation :

Proposition 3.6. Soient X est une variable aléatoire réelle discrète, $a, b \in \mathbb{R}$, alors

$$Var(aX + b) = a^2\ Var(X).$$

En particulier, $Var(X + b) = Var(X)$.

Définition 3.11. Soient X une variable discrète, p_X sa fonction de masse et $r \in \mathbb{N}^*$.

i. On appelle moment centré d'ordre r, la quantité :

$$\mu_r = \mathbb{E}\Big[\Big(X - \mathbb{E}(X)\Big)^r\Big] = \sum_{k \in X(\Omega)} \Big(k - \mathbb{E}(X)\Big)^r p_X(k).$$

ii. On appelle moment d'ordre r, la quantité :

$$\nu_r = \mathbb{E}\Big(X^r\Big) = \sum_{k\in X(\Omega)} k^r p_X(k).$$

Remarque 3.10. Soit X une variable discrète.

i. Si X est centrée, alors ses moments centrés et non centrés coïncident.

ii. Le moment centré d'ordre 2 de X est exactement sa variance.

iii. Le moments et les moments centrés d'une v. a. r discrète, appartiennent en général à $\mathbb{R}\cup\{-\infty,+\infty\}$.

3.3 Variables aléatoires absolument continues

Définition 3.12. Soient $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire réelle, F_X sa fonction de répartition. On dit que X est une variable absolument continue s'il existe une fonction réelle notée f_X qui vérifie

$$\forall x \in \mathbb{R}, \qquad F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{3.3}$$

La fonction f_X est appelée la fonction de densité de X.

Un critère simple pour vérifier si une v. a. r est absolument continue est donné par le résultat suivant :

Proposition 3.7. Soient $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire réelle, F_X sa fonction de répartition. Si F_X est une foncion dérivable sur $\mathbb{R}$ (sauf peut être en un nombre fini de points), alors X est une variable absolument continue. De plus, si on pose

$$\mathcal{D} = \{x\in\mathbb{R},\ t \to F_X(t) \text{ est dérivable au point } t = x\},$$

alors la fonction de densité f_X peut être définie par

$$f_X(x) = \begin{cases} F_X'(x) & \text{si } x\in\mathcal{D} \\ 0 & \text{sinon} \end{cases}$$

Comme dans le cas d'une variable aléatoire discrète, on peut se demander sous quelles conditions une fonction f est la fonction de densité d'une certaine variable aléatoire absolument continue ? Le résultat suivant fournit une réponse à cette question :

Proposition 3.8. Soit g une fonction réelle vérifiant :

i. Pour tout $x \in \mathbb{R}$, $g(x) \geq 0$.

ii. $\int_{\mathbb{R}} g(x)dx = 1$.

Alors il existe un espace de probabilité $(\Omega, \mathcal{F}, I\!\!P)$, une variable aléatoire X définie sur cet espace telle que g soit la fonction de densité de X, i. e. $f_X(x) = g(x)$, pour tout $x \in \mathbb{R}$.

Exemple 3.6. Soient $c \in \mathbb{R}$, f la fonction définie par

$$f(x) = \begin{cases} 2ce^{-\frac{x}{3}} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Sous quelle condition la fonction f est la fonction de densité d'une variable aléatoire réelle ?

<u>SOLUTION</u>:

Pour que f soit la fonction de densité d'une variable aléatoire il faut quelle soit positive, ainsi $c \geq 0$, de plus on doit avoir $\int_{\mathbb{R}} f(x)dx = 1$, par conséquent

$$1 = \int_{\mathbb{R}} f(x)dx = 2c \int_0^{+\infty} e^{-\frac{x}{3}} dx = 6c.$$

Ainsi, f est la fonction de densité d'une variable aléatoire si et seulement si $c = 1/6$.

Contrairement à ce qui ce passe dans le cas discret, pour une variable aléatoire absolument continue, la probabilité attachée à un point est toujours nulle, en effet

Proposition 3.9. Soit X une variable aléatoire <u>absolument continue</u>, alors

$$\forall x \in \mathbb{R}, \qquad I\!\!P(X = x) = 0. \tag{3.4}$$

PREUVE

Soient $\varepsilon > 0, x \in \mathbb{R}$ et soient F_X la fonction de répartition de X, f_X sa fonction de densité. Remarquons tout d'abord que

$$I\!\!P(X = x) = \lim_{\varepsilon \to 0} I\!\!P\Big(x - \varepsilon < X \leq x + \varepsilon\Big).$$

Mais par le théorème 3.2, on a

$$I\!\!P\Big(x - \varepsilon < X \leq x + \varepsilon\Big) = F_X(x + \varepsilon) - F_X(x - \varepsilon),$$

ainsi

$$\begin{array}{rcl} I\!\!P(X = x) & = & \lim\limits_{\varepsilon \to 0} \Big[F_X(x + \varepsilon) - F_X(x - \varepsilon)\Big] \\ & = & \lim\limits_{\varepsilon \to 0} \displaystyle\int_{x-\varepsilon}^{x+\varepsilon} f_X(t)dt \\ & = & 0. \end{array}$$

Remarque 3.11. On obtient comme conséquence immédiate de cette proposition : pour une variable aléatoire absolument continue X, pour $a, b \in \mathbb{R}$ avec $a < b$, les quantités $I\!\!P\Big(a < X \leq b\Big)$, $I\!\!P\Big(a \leq X \leq b\Big), I\!\!P\Big(a < X < b\Big)$, $I\!\!P\Big(a \leq X < b\Big)$ sont toute égales à $F_X(b) - F_X(a)$.

Proposition 3.10. Soient X une variable aléatoire absolument continue, f_X sa fonction de densité et $a, b \in \mathbb{R}$ avec $a < b$. On a

$$I\!\!P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt.$$

3.3.1 Espérance

Définition 3.13. On appelle espérance mathématique ou moyenne d'une variable aléatoire réelle absolument continue X, la quanité notée $\mathbb{E}(X)$ et définie par

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

où f_X est la fonction de densité de X.

Remarque 3.12. i. Il faut remarquer que l'espérance d'une v.a. r absolument continue est définie par une intégrale sur $\mathbb{R}$ tout entier c'est pour cette raison qu'elle appartient à $\mathbb{R} \cup \{-\infty, +\infty\}$ en général.

ii. Lorsque $\mathbb{E}(X) = 0$, on dit que la variable aléatoire X est centrée.

Proposition 3.11. Soient X une variable aléatoire réelle absolument continue, $g: \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue, posons $Y = g(X)$, alors Y est encore une v. a. r absolument continue et

$$\mathbb{E}(Y) = \mathbb{E}(g(X)) = \int_{\mathbb{R}} g(x) f_X(x) dx.$$

Définition 3.14. Soient $p \geq 1,\ X$ une variable aléatoire réelle absolument continue.

– On dit que X est intégrable si $\mathbb{E}(|X|) < \infty$.

– On dit que X est dans L^p si $\mathbb{E}(|X|^p) < \infty$.

Comme pour le cas discret, l'espérance dans le cas absolument continu est aussi linéaire, ceci est dû à la linéarité de l'intégrale, en effet

Proposition 3.12. Soient X, Y deux variables aléatoire absolument continues, $\alpha,\ \beta \in \mathbb{R}$, alors

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

ii. Si $\mathbb{E}(|X|) = 0$ alors $I\!P(X = 0) = 1$.

3.3.2 Variance et Ecart-type

Définition 3.15. Soit X une variable aléatoire réelle absolument continue

– On appelle variance de X, la quantité notée $Var(X)$ et définie par

$$Var(X) = \mathbb{E}\left[\Big(X - \mathbb{E}(X)\Big)^2\right] = \int_R \Big(x - \mathbb{E}(X)\Big)^2 f_X(x) dx.$$

– On appelle écart-type de X, la quantité notée $\sigma(X)$ et définie par

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

Remarque 3.13. La variance d'une v. a. r absolument continue est définie par l'intégrale sur $\mathbb{R}$ d'une fonction positive, par conséquent $Var(X) \in \mathbb{R}^+ \cup \{+\infty\}$ en général.

En développant le carré dans la définition de la variance, on obtient une nouvelle formule plus utile dans les calculs, en effet :

Proposition 3.13. Si X est une variable aléatoire réelle absolument continue, alors

$$Var(X) = \mathbb{E}(X^2) - \Big(\mathbb{E}(X)\Big)^2 = \int_R x^2 f_X(x)dx \ - \ \Big(\mathbb{E}(X)\Big)^2.$$

Contrairement à l'espérance mathématique, la variance n'est pas linéaire, elle est cependant invariante par translation :

Proposition 3.14. Soient X est une variable aléatoire réelle absolument continue, $a, b \in \mathbb{R}$, alors

$$Var(aX + b) = a^2 \ Var(X).$$

Définition 3.16. Soient X une variable aléatoire réelle absolument continue, $r \in \mathbb{N}^*$.

i. On appelle <u>moment centré d'ordre r</u>, la quantité :

$$\mu_r = \mathbb{E}\Big[\Big(X - \mathbb{E}(X)\Big)^r\Big] = \int_{\mathbb{R}} \Big(x - \mathbb{E}(X)\Big)^r f_X(x)dx.$$

ii. On appelle <u>moment d'ordre r</u>, la quantité :

$$\nu_r = \mathbb{E}\Big(X^r\Big) = \int_{\mathbb{R}} x^r f_X(x)dx.$$

Remarque 3.14. Soit X une v. a. r absolument continue.

i. Si X est centrée, alors ses moments centrés et non centrés coïncident.

ii. Le moment centré d'ordre 2 de X est exactement sa variance.

iii. Les moments et le moments centrés d'une v. a. r absolument continue, appartiennent en générale à $\mathbb{R} \cup \{-\infty, +\infty\}$.

3.4 Fonction génératrice des moments

Pour une variable aléatoire X, posons

$$\mathcal{H} = \{t \in \mathbb{R}/ \ \mathbb{E}\Big(e^{tX}\Big) < \infty\}.$$

L'ensemble $\mathcal{H}$ est tout simplement le domaine de définition de la fonction $t \mapsto \mathbb{E}\Big(e^{tX}\Big)$, il faut remarquer que cet ensemble ne peut pas être vide puisque $0 \in \mathcal{H}$.

Définition 3.17. Soit X une variable aléatoire réelle, la fonction M_X définie par

$$M_X(t) = \mathbb{E}\Big(e^{tX}\Big), \qquad t \in \mathcal{H} \tag{3.5}$$

est appelée la fonction génératrice des moments de la variable X.

Exemple 3.7. Calculons la fonction génératrice des moments pour quelques variables aléatoires :

i. Soit X la v. a. r qui prends les valeurs $+1$ et -1 avec une probabilité $\frac{1}{2}$, alors un petit calcul donne :

$$\mathcal{H} = \mathbb{R} \quad \text{ et } \ M_X(t) = \cosh(t) = \frac{e^t + e^{-t}}{2}.$$

ii. Soient $\lambda > 0$, Y la v. a. r absolument continue de fonction de densité :

$$f_Y(x) = \lambda e^{-\lambda x} \mathbb{1}_{]0,+\infty[}(x),$$

alors

$$M_Y(t) = \mathbb{E}(e^{tY}) = \int_{\mathbb{R}} e^{tx} f_Y(x) dx = \lambda \int_0^{\infty} e^{(t-\lambda)x} dx.$$

Remarquons d'abord que $M_Y(\lambda) = +\infty$. Supposons donc que $t \neq \lambda$. On a

$$\begin{aligned} M_Y(t) &= \lambda \int_0^{\infty} e^{(t-\lambda)x} dx \\ &= \frac{\lambda}{t-\lambda} [e^{(t-\lambda)x}]_0^{\infty}. \end{aligned}$$

Par suite

$$M_Y(t) = \begin{cases} \frac{\lambda}{\lambda-t} & \text{si } t < \lambda \\ +\infty & \text{si } t \geq \lambda \end{cases}$$

Ceci implique que $\mathcal{H} =]-\infty, \lambda[$ et

$$\forall t \in \mathcal{H}, \quad M_Y(t) = \frac{\lambda}{\lambda - t}.$$

La fonction génératrice des moments permet d'identifier la loi d'une variable aléatoire, en effet

Théorème 3.5. Soient X, Y deux variables aléatoires réelles, M_X, M_Y leur fonctions génératrices, alors

$$\Big[X \text{ et } Y \text{ ont même loi}\Big] \iff \Big[M_X \equiv M_Y\Big].$$

Une autre propriété importante de la fonction génératrice des moments est qu'elle permet de calculer les moments d'une v. a. r lorsqu'un calcul direct s'avère difficile, en effet :

Théorème 3.6. Soient X une variables aléatoire réelle, M_X sa fonction génératrice des moments. Si la fonction $t \to M_X(t)$ est définie sur un intervalle <u>ouvert et contenant 0</u>, alors elle admet le développement en série entière suivant :

$$M_X(t) = \sum_{n \in \mathbb{N}} \frac{\mathbb{E}(X^n)}{n!} t^n,$$

pour tout t dans un voisinage de 0. En particulier, si $n \geq 1$, alors

$$\mathbb{E}(X^n) = M_X^{(n)}(0),$$

où $M_X^{(n)}$ désigne la dérivée d'ordre n de M_X.

Exemple 3.8. Calculons la moyenne et la variance pour les variables X et Y de l'exemple 3.7. On a

$$\forall t \in \mathbb{R}, \ \forall u \in]-\infty, \lambda[, \quad M_X(t) = \cosh(t) \ \text{ et } \ M_Y(u) = \frac{\lambda}{\lambda - u}.$$

Ainsi, les deux fonction génératrices vérifient les conditions du théorème 3.6, d'où

$$\mathbb{E}(X) = M_X'(0) = \sinh(0) = 0, \qquad \mathbb{E}(X^2) = M_X''(0) = \cosh(0) = 1.$$

et

$$\mathbb{E}(Y) = M_Y'(0) = \frac{1}{\lambda}, \qquad \mathbb{E}(Y^2) = M_X''(0) = \frac{2}{\lambda^2}$$

Par suite

$$Var(X) = 1 \qquad \text{et} \qquad Var(Y) = \frac{1}{\lambda^2}.$$

3.5 Fonction caractéristique

Dans ce paragraphe $\mathbb{C}$ désigne l'ensemble des nombres complexes, i le nombre complexe vérifiant : $i^2 = -1$.

Définition 3.18. Soit $X : \Omega \longrightarrow \mathbb{C}$ une application. On dit que X est une variable aléatoire complexe si ses parties réelle et imaginaire sont des variables aléatoires réelles.

Définition 3.19. Soit X une variable aléatoire réelle, la fonction $\varphi_X : \mathbb{R} \longrightarrow \mathbb{C}$ définie par

$$\varphi_X(t) = \mathbb{E}\Big(e^{itX}\Big), \qquad t \in \mathbb{R} \tag{3.6}$$

est appelée la fonction caractéristique de la variable X.

Contrairement à la fonction génératrice qui peut être infinie, la fonction caractéristique est toujours finie, mieux elle est bornée, en effet

Proposition 3.15. Soient X une v. a. r, φ_X sa fonction caractérisque, alors

1. $\varphi_X(0) = 1$ et $\sup_{t\in\mathbb{R}} |\varphi_X(t)| \leq 1$.
2. La fonction $t \mapsto \varphi_X(t)$ est uniformément continue.
3. X est une v. a. r symétrique (i. e $\mathcal{L}(X) = \mathcal{L}(-X)$), si et seulement si φ_X est une fonction réelle paire

4. Soient $a, b \in \mathbb{R}$ et posons $Y = aX + b$, alors

$$\forall t \in \mathbb{R}, \qquad \varphi_Y(t) = e^{ibt}\varphi_X(at).$$

Calculons a présent la fonction caractéristique pour quelques variables aléatoires :

Exemple 3.9. Considérons les variables suivantes :

i. Soit X la v. a. r qui prends les valeurs $+1$ et -1 avec une probabilité $\frac{1}{2}$, alors un petit calcul donne :

$$\varphi_X(t) = \mathbb{E}\Big(e^{itX}\Big) = \frac{e^{it} + e^{-it}}{2} = \cos(t).$$

ii. Soit Y la v. a. r absolument continue de fonction de densité :

$$f_Y(x) = \mathbb{1}_{[0,1]}(x).$$

Pour $t = 0$, on a $\varphi_Y(0) = 1$. Si $t \neq 0$, alors

$$\varphi_Y(t) = \mathbb{E}\Big(e^{itY}\Big) = \int_0^1 e^{ity}dy = \frac{e^{it} - 1}{it}.$$

Donc

$$\varphi_Y(t) = \begin{cases} \frac{e^{it}-1}{it} & \text{si } t \neq 0 \\ 1 & \text{si} \quad t = 0. \end{cases}$$

La fonction génératrice, permet d'identifier la loi d'une variable aléatoire en effet :

Théorème 3.7. Soient X, Y deux variables aléatoires, φ_X, φ_Y leur fonctions caractéristiques, alors

$$\Big[X \text{ et } Y \text{ ont même loi}\Big] \iff \Big[\varphi_X \equiv \varphi_Y\Big].$$

Comme la fonction ganératrice des moments, la fonction caractéristique permet de calculer les moments d'une variable réelle lorsque le caclul directe est difficile.

Proposition 3.16. Soient X deux variable aléatoire, φ_X sa fonction caractéristique et $n \geq 1$. Supposons que $\mathbb{E}(X^n) < \infty$, alors la fonction φ_X est n fois dérivable, de plus

$$\forall 1 \leq k \leq n, \qquad \mathbb{E}(X^k) = (-i)^k \varphi^{(k)}(0),$$

où $\varphi^{(k)}$ est la dérivée d'ordre k de φ.

Exemple 3.10. Soient X, Y les variables aléatoires de l'exemple 3.9. On rappelle que

$$\varphi_X(t) = \cos(t), \quad \text{et} \quad \varphi_Y(u) = \frac{e^{iu} - 1}{iu},$$

pour $t \in \mathbb{R}$, $u \in \mathbb{R}^*$ et $\varphi_Y(0) = 1$. Maintenant il faut remarquer que les deux variables sont bornées, donc elle admettent des moments de tout ordre, ainsi on peut appliquer le théorème 3.7. On a

$$\varphi_X^{'}(0) = 0 \quad \text{et} \quad \varphi_Y^{'}(0) = \frac{i}{2}$$

Par suite

$$\mathbb{E}(X) = -i\varphi_X^{'}(0) = 0 \quad \text{et} \quad \mathbb{E}(Y) = -i\varphi_Y^{'}(0) = \frac{1}{2}.$$

De même

$$\varphi_X^{''}(0) = -1 \quad \text{et} \quad \varphi_Y^{''}(0) = -\frac{1}{3}$$

Par suite

$$\mathbb{E}(X^2) = -\varphi_X^{''}(0) = 1 \quad \text{et} \quad \mathbb{E}(Y^2) = -\varphi_Y^{''}(0) = \frac{1}{3}$$

Par conséquent

$$Var(X) = 1 \quad \text{et} \quad Var(Y) = \frac{1}{12}.$$

3.6 Trois inégalités utiles

Dans ce paragraphe on donne trois inégalités qu'on utilisera dans les chapitres suivants, ces inégalités sont valables pour les variables aléatoires discrètes et les variables aléatoires absolument continues, leur interêt réside dans le fait qu'elle ne font pas appel à la loi de la variable alétoire considérée. Nous commençons par l'inégalité de Jensen.

Définition 3.20. Soit I un intervalle de $\mathbb{R}$, $f : I \longrightarrow \mathbb{R}$ une fonction. On dit que f est convexe sur I si

$$\forall x, y \in I,\ \forall \lambda \in [0,1], \quad f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Exemple 3.11. La fonction $x \longrightarrow |x|$ est convexe sur $\mathbb{R}$.

En général, il est difficile de vérifier qu'une fonction donnée est convexe en utilisant juste la définition, cependant si cette fonction est deux fois dérivable, alors

Proposition 3.17. Soit I un intervalle ouvert de $\mathbb{R}$, $f : I \longrightarrow \mathbb{R}$ une fonction deux fois dérivable sur I, alors

$$f \text{ convexe sur } I \Longleftrightarrow \ \forall x \in I,\ f^{''}(x) \geq 0.$$

Comme conséquence de ce résultat, on a

Corollaire 3.1. Soient $p \geq 1, n \in \mathbb{N}^*$ et $\lambda \in \mathbb{R}$, alors

i. $x \longrightarrow x^p$ est convexe sur $\mathbb{R}^+$.

ii. $x \longrightarrow x^{2n}$ est convexe sur $\mathbb{R}$.

i. $x \longrightarrow e^{\lambda x}$ est convexe sur $\mathbb{R}$.

A présent on peut énnoncer l'inégalité de Jensen :

Proposition 3.18. Soient $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe, X une variable aléatoire réelle. Supposons que $\mathbb{E}\Big(|f(X)|\Big) < \infty$, alors on a l'inégalité de Jensen :

$$f\Big(\mathbb{E}(X)\Big) \leq \mathbb{E}\Big(f(X)\Big) \tag{3.7}$$

En appliquant ce résultat à des fonctions convexes particulières, on a

Corollaire 3.2. Soient X une variable aléatoire réelle, $n \in \mathbb{N}^*$ et $\lambda \in \mathbb{R}$.

i. Si $\mathbb{E}\Big(|X|\Big) < \infty$, alors
$$|\mathbb{E}\Big(X\Big)| \leq \mathbb{E}\Big(|X|\Big).$$

ii. Si $\mathbb{E}\Big(X^{2n}\Big) < \infty$, alors
$$\Big(\mathbb{E}(X)\Big)^{2n} \leq \mathbb{E}\Big(X^{2n}\Big).$$

iii. Si X est positive et si $\mathbb{E}\Big(X^n\Big) < \infty$, alors
$$\Big(\mathbb{E}(X)\Big)^{n} \leq \mathbb{E}\Big(X^{n}\Big).$$

iv. Si $\mathbb{E}\Big(e^{\lambda X}\Big) < \infty$, alors
$$e^{\lambda \mathbb{E}(X)} \leq \mathbb{E}\Big(e^{\lambda X}\Big).$$

La deuxième inégalité est utile pour estimer la quantité $I\!P(|X| > \lambda)$ lorsqu'on connait pas la loi de la variable aléatoire X, elle est connue sous le nom inégalité de Markov :

Proposition 3.19. Soient $\lambda > 0$, X une variable aléatoire réelle, avec $\mathbb{E}\Big(|X|\Big) < \infty$, alors
$$I\!P\Big(|X| > \lambda\Big) \leq \frac{\mathbb{E}(|X|)}{\lambda}. \tag{3.8}$$

La dernière inégalité est une conséquence de l'inégalité de Markov, elle appelée inéégalité de Tchebycheff :

Proposition 3.20. Soient $\lambda > 0$, X une variable aléatoire réelle, avec $\mathbb{E}\Big(X^2\Big) < \infty$, alors
$$I\!P\Big(|X - \mathbb{E}(X)| > \lambda\Big) \leq \frac{Var(X)}{\lambda^2}. \tag{3.9}$$

3.7 Exercices

Exercice 3.1. Soient Ω un espace échantillon, A et B deux sous ensembles de Ω avec $A \neq B$. Déterminer l'algèbre engendrée par la famille $\mathcal{C} = \{A, B\}$.

Exercice 3.2. Soient $(\Omega, \mathcal{F})$ un espace probabilsable et $X : \Omega \to \mathbb{R}$ une application, posons

$$\sigma_X = \{X^{-1}(A);\ A \subset \mathbb{R}\}.$$

Montrer que σ_X est une tribu sur Ω, elle est appelée la tribu engendrée par X.

Exercice 3.3. Soient Ω un espace échantillon, $\mathcal{C} \subset \mathcal{P}(\Omega)$. Posons

$$\mathcal{T}(\mathcal{C}) = \{\mathcal{T} \subset \mathcal{P}(\Omega),\ \mathcal{T} \text{ tribu },\ \mathcal{C} \subset \mathcal{T}\}$$

1. Montrer que $\mathcal{T}(\mathcal{C})$ n'est pas vide.
2. Posons

$$\sigma(\mathcal{C}) = \bigcap_{\mathcal{T} \in \mathcal{T}(\mathcal{C})} \mathcal{T}.$$

Montrer que $\sigma(\mathcal{C})$ est la plus petite tribu contenant $\mathcal{C}$. ($\sigma(\mathcal{C})$ *est appelée la tribu engendrée par* $\mathcal{C}$).

Exercice 3.4. Soient $a, b \in \mathbb{R}$ avec $a < b$. Montrer que les ensembles suivants sont des boréliens :

$$[a, +\infty[,\ ]a, +\infty[,\ ]-\infty, a[,\ [a, b],\ [a, b[,\ ]a, b[,\ ]a, b],\ \{a\}.$$

Exercice 3.5. Soit p la fonction définie par

$$p(x) = c\frac{(x-1)}{n} \qquad \text{si } \ x \in \{1, 2, ..., n\}.$$

– Déterminer la valeur de c pour que p soit la fonction de masse d'une variable aléatoire X.
– Déterminer dans ce cas F_X.
– Calculer $I\!P(X \leq 3)$, $I\!P(1 < X \leq 5)$ et $I\!P(X > n-2)$.
– Calculer la moyenne de X.

Exercice 3.6. Soit X une variable aléatoire discrète de fonction de masse :

$$p_X(k) = \frac{e^{-2}2^k}{4k!}(1+ck), \qquad k \in \mathbb{N}.$$

- Déterminer c.
- Soient $Y,\ Z$ deux variable aléatoire de Poissons de paramètre 2. Montrer que
$$\mathbb{P}(X=k) = \frac{1}{4}\mathbb{P}(Y=k) + \frac{3}{4}\mathbb{P}(Z=k-1).$$
- En déduire $\mathbb{E}(X)$.

Exercice 3.7. Soit f la fonction définie par

$$f(x) = cxe^{-\frac{x}{2}}\mathbb{1}_{[0,+\infty[}(x).$$

- Déterminer la valeur de c pour que f soit une fonction de densité d'une variable aléatoire X.
- Calculer la fonction de répartition de X.
- Calculer la moyenne et la variance de X.

Exercice 3.8. Soit X une variable aléatoire réelle absolument continue de fonction de densité

$$f_X(x) = ke^{-|x-5|}, \qquad x \in \mathbb{R}.$$

a. Déterminer la constante k.

b. Calculer F_X, la fonction de répartition de X ainsi que la moyenne et la variance de X.

Exercice 3.9. La durée de vie d'une machine est modélisée par une variable aléatoire X. D'après des études statistiques, on sait que ce type de machine fonctionne encore après 9 ans avec une probabilité 0.1. On propose

$$f(x) = \frac{a}{(x+1)^b}\mathbb{1}_{[0,+\infty[}(x).$$

Comme fonction de densité pour cette variable aléatoire.

- Déterminer les constantes a et b.

– Calculer la durée moyenne de fonctionnement pour ce type de machine.

– Quelle est la probabilité que cette machine fonctionne encore après 12 ans.

Exercice 3.10. Soit X une variable aléatoire de fonction de densité f donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

Posons

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{itx-\frac{x^2}{2}} dx.$$

1. Calculer φ' en intégrant sous le signe intégrale.
2. Montrer que la fonction φ vérifie une equation différentielle à déterminer.
3. En déduire la fonction caractéristique de la variable X.

Chapitre 4

Lois usuelles

Le but de ce chapitre est de donner les principales caractéristiques des variables aléatoires les plus utilisées en calcul des probabilités, nous commençons par les variables aléatoires discrètes. toutes les variables aléatoires de ce chapitre sont définies sur un espace de probabilités $(\Omega, \mathcal{F}, I\!P)$.

4.1 Lois discrètes

4.1.1 Loi de Dirac

Définition 4.1. Soient $a \in \mathbb{R}$, X une variable aléatoire réelle. On dit que X suit une loi de Dirac concentrée en a si $I\!P(X = a) = 1$. (On note $X \hookrightarrow \delta_a$).

Remarque 4.1. Il n'est pas difficile de voir que si $X \hookrightarrow \delta_a$, alors sa fonction de masse est donnée par

$$p_X(x) = \begin{cases} 1 & \text{si } x = a \\ 0 & \text{si } x \neq a \end{cases}$$

Par conséquent, $\mathbb{E}(X) = a$ et $Var(X) = 0 = \sigma_X^2$.

Proposition 4.1. Soient $X \hookrightarrow \delta_a$, M_X, sa fonction génératrice des moments et φ_X sa fonction caractéristique, alors

$$M_X(t) = e^{at}, \qquad \varphi_X(t) = e^{iat},$$

pour tout $t \in \mathbb{R}$.

4.1.2 Loi de Bernoulli

Définition 4.2. Soient $0 < p < 1$, X une variable aléatoire réelle. On dit que X suit une loi de bernoulli de paramètre p (on note $X \hookrightarrow \mathcal{B}(p)$) si X est à valeurs dans $\{0, 1\}$ avec

$$P(X = 1) = p = 1 - P(X = 0).$$

Remarque 4.2. La variable de Bernoulli est souvent utilisée pour modéliser une experience aléatoire qui n'a que deux issues. En général, le chiffre 1 est atribué au "succès" alors que 0 est attribué à "l'echec". Dans une expérience aléatoire de Bernoulli, le paramètre p est donc la probabilité du "succès".

Il n'est pas difficile de montrer

Proposition 4.2. Soit $X \hookrightarrow \mathcal{B}(p)$, alors

1. La fonction de masse de X est donnée par

$$p_X(k) = \begin{cases} p^k(1-p)^{1-k} & \text{si } k \in \{0, 1\} \\ 0 & \text{sinon} \end{cases}$$

2. $\mathbb{E}(X) = p$, $Var(X) = p(1-p)$ et $\sigma_X = \sqrt{p(1-p)}$.

Proposition 4.3. Soient $X \hookrightarrow \mathcal{B}(p)$, M_X, sa fonction génératrice des moments et φ_X sa fonction caractéristique, alors

$$M_X(t) = 1 - p + pe^t, \qquad \varphi_X(t) = 1 - p + pe^{it},$$

pour tout $t \in \mathbb{R}$.

4.1.3 Loi binomiale

Définition 4.3. Soient $n \geq 1$ et $0 < p < 1$. On répète une experience de Bernoulli n fois de manière indépendantes et on note par X la variable aléatoitre qui compte le nombre de fois où on a obtenu le chiffre 1. On dit que la variable X ainsi définie suit une loi binomiale de paramètres n et p. (on note $X \hookrightarrow \mathcal{B}(n, p)$)

Remarque 4.3. d'après cette définition, on peut affirmer que

- Une variable aléatoire qui suit une loi binomiale de paramètres n et p est tout simplement une variable de comptage, elle compte le nombre de "succès" obtenus lorsqu'on répète la même expérience plusieurs fois.
- Le paramètre n est un entier positif, il désigne le nombre de répétitions, alors que p est la probabilité du "succès"
- Une variable binomiale $\mathcal{B}(n, p)$, est à valeurs dans l'ensemble $\{0, 1, ..., n\}$.

Proposition 4.4. Soit $X \hookrightarrow \mathcal{B}(n, p)$, alors

1. La fonction de masse de X est donnée par

$$p_X(k) = \begin{cases} C_n^k p^k (1-p)^{n-k} & \text{si } k \in \{0, 1, .., n\} \\ 0 & \text{sinon} \end{cases}$$

2. $\mathbb{E}(X) = np,\ Var(X) = np(1-p)$ et $\sigma_X = \sqrt{np(1-p)}$.

Exemple 4.1. Dans un exercice militaire, un soldat a le droit de tirer sur une cible mobile 10 fois, si la probabilité d'atteindre cette cible est 0.7, quelle est la probabilité que ce soldat atteigne la cible au moins 2 fois.

<u>SOLUTION</u>:

Cette expérience aléatoire consiste à répéter la même expérience (tirer sur une cible) 10 fois de suite. c'est donc une expérience binomiale. Soit X la variable aléatoire qui modélise cette expérience, on a $X \hookrightarrow \mathcal{B}(10, 0.7)$ et on cherche $I\!P\Big(X \geq 2\Big)$. Donc

$$\begin{aligned} I\!P\Big(X \geq 2\Big) &= 1 - I\!P\Big(X < 2\Big) \\ &= 1 - \Big(I\!P\Big(X = 0\Big) + I\!P\Big(X = 1\Big)\Big) \\ &= 1 - \Big(C_{10}^0 (1-0.7)^{10} + C_{10}^1 (0.7)(1-0.7)^9\Big) \\ &= 0.856. \end{aligned}$$

Proposition 4.5. Soient $X \hookrightarrow \mathcal{B}(n, p),\ M_X$ sa fonction génératrice des moments et φ_X sa fonction caractéristique, alors

$$M_X(t) = (1 - p + pe^t)^n, \qquad \varphi_X(t) = (1 - p + pe^{it})^n,$$

pour tout $t \in \mathbb{R}$.

4.1.4 Loi de Poisson

Définition 4.4. Soient $\lambda > 0$. On dit que la variable X suit une loi de Poisson de paramètre λ (on note $X \hookrightarrow \mathcal{P}(\lambda)$), si sa fonction de masse est donnée par

$$p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k \in \mathbb{N}.$$

Remarque 4.4. On rappelle que si $x \in \mathbb{R}$, alors

$$e^x = \sum_{k\in\mathbb{N}} \frac{x^k}{k!}.$$

Ainsi p_X est bien une fonction de masse.

Proposition 4.6. Soit $X \hookrightarrow \mathcal{P}(\lambda)$, alors

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

Exemple 4.2. Dans la gare ferroviaire d'Oran, un guichet vend en moyenne 2 billets chaque demi-heure pour le trajet Oran-Alger. Quelle est la probabilité que ce guichet vende 6 billets pour ce trajet dans un intervalle de 2 heures ?
SOLUTION:

Soit X la v. a. r qui modélise le nombre de billets vendu par ce guichet pendant 2 heures. Il faut remarquer que pour une loi de Poisson, la moyenne est exactement le paramètre, ainsi pour 2 heures la moyenne est $\mathbb{E}(X) = 8$. Par suite on cherche $\mathbb{P}\Big(X = 6\Big)$. on a

$$\mathbb{P}\Big(X = 6\Big) = e^{-8}\frac{8^6}{6!} = 0.106$$

Proposition 4.7. Soient $X \hookrightarrow \mathcal{P}(\lambda)$, M_X sa fonction génératrice des moments et φ_X sa fonction caractéristique, alors

$$M_X(t) = e^{\lambda(e^t-1)}, \qquad \varphi_X(t) = e^{\lambda(e^{it}-1)},$$

pour tout $t \in \mathbb{R}$.

4.1.5 Loi géométrique

Définition 4.5. Soient $0 < p < 1$. On dit que la variable X suit une loi géométrique de paramètre p (on note $X \hookrightarrow \mathcal{G}(p)$), si sa fonction de masse est donnée par

$$p_X(k) = p(1-p)^{k-1}, \qquad k \in \mathbb{N}^*.$$

Proposition 4.8. Soit $X \hookrightarrow \mathcal{G}(p)$, alors

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

Proposition 4.9. Soient $X \hookrightarrow \mathcal{G}(p)$, M_X, sa fonction génératrice des moments et φ_X sa fonction caractéristique, alors

$$M_X(t) = \frac{pe^t}{1-(1-p)e^t} \qquad \text{et} \qquad \varphi_X(u) = \frac{pe^{iu}}{1-(1-p)e^{iu}},$$

pour $t \in]-\infty, -\ln(1-p)[$ et $u \in \mathbb{R}$.

4.1.6 Loi hypergéométrique

Définition 4.6. Soit une urne contenant N boules dont une proportion p de boules blanches. On prélève de cette urne un échantillon (sans remise) de n boules. On note par X la variable aléatoire qui donne le nombre de boules blanches de l'échantillon. On dit que X suit une loi hypergéométrique de paramètres N, n, p et on note $X \hookrightarrow \mathcal{H}(N, n, p)$.

D'après le chapitre 1, on a

Proposition 4.10. Soit $X \hookrightarrow \mathcal{H}(N, n, p)$, alors

$$p_X(k) = \begin{cases} \frac{C_{Np}^k C_{qN}^{n-x}}{C_N^n} & \text{si } \max(0, n-Nq) \leq k \leq \min(n, Np) \\ 0 & \text{sinon} \end{cases}$$

où $q = 1-p$.

On a d'autre part

Proposition 4.11. Soit $X \hookrightarrow \mathcal{H}(N, n, p)$, alors

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

4.2 Lois absolument continue

4.2.1 Loi uniforme

Définition 4.7. Soient $a, b \in \mathbb{R}$ avec $a < b$. On dit qu'une variable aléatoire X suit une loi uniforme sur $[a, b]$ si sa fonction de densité f est donnée par

$$f(x) = \frac{1}{b-a} \ \mathbb{1}_{[a,b]}(x)$$

On note $X \hookrightarrow \mathcal{U}_{[a,b]}$.

Il n'est pas difficile de montrer

Proposition 4.12. Soit $X \hookrightarrow \mathcal{U}_{[a,b]}$, alors

$$\mathbb{E}(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12} \quad \text{et} \quad \sigma_X = \frac{b-a}{2\sqrt{3}}.$$

De même, on a

Proposition 4.13. Soit $X \hookrightarrow \mathcal{U}_{[a,b]}$, alors

$$M_X(t) = \begin{cases} \frac{e^{bt}-e^{at}}{(b-a)t} & \text{si } t \neq 0 \\ 1 & \text{sinon} \end{cases}$$

et

$$\varphi_X(t) = \begin{cases} \frac{e^{ibt}-e^{iat}}{i(b-a)t} & \text{si } t \neq 0 \\ 1 & \text{sinon} \end{cases}$$

4.2.2 Loi exponentielle

Définition 4.8. Soit $\lambda > 0$. On dit qu'une variable aléatoire X suit une loi exponentielle de paramètre λ si sa fonction de densité f est donnée par

$$f(x) = \lambda e^{-\lambda x} \, \mathbb{1}_{[0,+\infty[}(x)$$

On note $X \hookrightarrow \mathcal{E}(\lambda)$.

Il n'est pas difficile de montrer

Proposition 4.14. Soit $X \hookrightarrow \mathcal{E}(\lambda)$, alors

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

De même, on a

Proposition 4.15. Soit $X \hookrightarrow \mathcal{E}(\lambda)$, alors

$$M_X(t) = \frac{\lambda}{\lambda - t} \qquad \text{et} \qquad \varphi_X(u) = \frac{\lambda}{\lambda - iu},$$

pour $t \in]-\infty, \lambda[$ et $u \in \mathbb{R}$.

4.2.3 Loi gamma

Définition 4.9. Soient $r > 0, \lambda > 0$. On définit la fonction Γ par

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

On dit qu'une variable aléatoire X suit une loi gamma de paramètres r, λ si sa fonction de densité f est donnée par

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \, \mathbb{1}_{]0,+\infty[}(x)$$

On note $X \hookrightarrow \Gamma(\lambda, r)$.

Proposition 4.16. Soit $X \hookrightarrow \Gamma(\lambda, r)$, alors

$$\mathbb{E}(X) = \frac{r}{\lambda}, \quad Var(X) = \frac{r^2}{\lambda^2} \quad \text{et} \quad \sigma_X = \frac{r}{\lambda}.$$

De même

Proposition 4.17. Soit $X \hookrightarrow \Gamma(\lambda, r)$, alors

$$M_X(t) = \Big(\frac{\lambda}{\lambda - t}\Big)^r, \qquad \varphi_X(u) = \Big(\frac{\lambda}{\lambda - iu}\Big)^r,$$

pour $t \in]-\infty, \lambda[$ et $u \in \mathbb{R}$.

4.2.4 Loi de Laplace

Définition 4.10. On dit qu'une variable aléatoire suit une loi de Laplace si sa fonction de densité est donnée par

$$f(x) = \frac{1}{2}e^{-|x|}, \qquad x \in \mathbb{R}.$$

On note $X \hookrightarrow \mathcal{L}$.

Proposition 4.18. Soit $X \hookrightarrow \mathcal{L}$, alors

$$\mathbb{E}(X) = 0, \quad Var(X) = 2 \quad \text{et} \quad \sigma_X = \sqrt{2}.$$

Proposition 4.19. Soit $X \hookrightarrow \mathcal{L}$, alors

$$M_X(t) = \frac{1}{1 - t^2}, \qquad \varphi_X(u) = \frac{1}{1 + u^2},$$

pour tout $t \in]-1, 1[$ et $u \in \mathbb{R}$.

4.2.5 Loi de Cauchy

Définition 4.11. On dit qu'une variable aléatoire X suit une loi de Cauchy si sa fonction de densité f est donnée par

$$f(x) = \frac{1}{\pi}\,\frac{1}{1+x^2}, \qquad x \in \mathbb{R}.$$

On note $X \hookrightarrow \mathcal{C}$.

Proposition 4.20. Soit $X \hookrightarrow \mathcal{C}$, alors

$$\varphi_X(t) = e^{-|t|},$$

pour tout $t \in \mathbb{R}$.

4.2.6 Loi normale centrée réduite

Définition 4.12. On dit qu'une variable aléatoire X suit une loi de normale centrée réduite (ou gaussienne centrée réduite) si sa fonction de densité f est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}, \qquad x \in \mathbb{R}.$$

On note $X \hookrightarrow \mathcal{N}(0,1)$.

Proposition 4.21. Soit $X \hookrightarrow \mathcal{N}(0,1)$, alors

$$\mathbb{E}(X) = 0, \quad Var(X) = 1 = \sigma_X.$$

Proposition 4.22. Soit $X \hookrightarrow \mathcal{N}(0,1)$, alors

$$M_X(t) = e^{\frac{t^2}{2}}, \qquad \varphi_X(t) = e^{-\frac{t^2}{2}},$$

pour tout $t \in \mathbb{R}$.

4.2.7 Loi normale

Définition 4.13. Soient $\mu \in \mathbb{R},\ \sigma > 0$. On dit qu'une variable aléatoire X suit une loi de normale ou gaussienne si sa fonction de densité f est donnée par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad x \in \mathbb{R}.$$

On note $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

Proposition 4.23. Soit $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors

$$\mathbb{E}(X) = \mu, \quad Var(X) = \sigma^2, \qquad \sigma_X = \sigma.$$

Proposition 4.24. Soit $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$, alors

$$M_X(t) = e^{\mu t + \frac{t^2}{2\sigma^2}}, \qquad \text{et} \qquad \varphi_X(t) = e^{i\mu t - \frac{t^2}{2\sigma^2}},$$

pour tout $t \in \mathbb{R}$.

4.3 Approximation de la loi binomiale

Soient $n \geq 1, 0 < p < 1$ et $X \hookrightarrow \mathcal{B}(n,p)$ et supposons que le produit np tends vers $\lambda > 0$ lorsque $n \to +\infty$ et $p \to 0$. Dans ce cas, pour n assez grand on peut remplacer p par $\frac{\lambda}{n}$. D'autre part si $k \in \{0, 1, ..., n\}$, alors

$$p_X(k) = I\!P(X = k) = C_n^k p^k (1-p)^{n-k},$$

ainsi pour tout $k \in \{0, 1, ..., n\}$:

$$\begin{aligned} p_X(k) &= \frac{n(n-1)\cdots(n-k+1)}{k!} p^k (1-p)^{n-k} \\ &= \frac{n^k(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})}{k!} (\frac{p}{1-p})^k (1-p)^n \\ &= \frac{(np)^k}{k!}(1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})(1-\frac{\lambda}{n})^n. \end{aligned}$$

Mais

$$\lim_{n\to+\infty} \frac{(np)^k}{k!} = \frac{\lambda^k}{k!} \quad \text{et} \quad \lim_{n\to+\infty} (1-\frac{\lambda}{n})^n = e^{-\lambda}.$$

Par conséquent

$$\lim_{n\to+\infty} p_X(k) = e^{-\lambda}\,\frac{\lambda^k}{k!}$$

qui est la fonction de masse d'une loi de Poisson de paramètre λ, on a démontré le résultat suivant :

Théorème 4.1. Soient $n \geq 1, 0 < p < 1$ et X une variable aléatoire de loi $\mathcal{B}(n,p)$. Supposons que $np \to \lambda$ lorsque $n \to +\infty$ et $p \to 0$. Alors

$$\mathcal{B}(n,p) \simeq \mathcal{P}(\lambda).$$

Afin d'illustrer ce résultat, considérons l'exemple suivant :

Exemple 4.3. Dans un pays le pourcentage des personne atteints d'une certaine maladie génétique est de 0.4%. Quelle est la probabilité de trouver au plus 2 personnes atteintes de cette maladie dans un village de 250 personnes ?

<u>Solution</u>:

Soit X la variable aléatoire qui donne le nombre de personnes atteinte de cette maladie dans ce village de 250 personnes. Il est clair qu'on cherche $\mathbb{P}(X \leq 2)$ et que $X \hookrightarrow \mathcal{B}(250, 0.004)$. Ainsi n et assez grand et p est petit, de plus le produit $np = 250 \times 0.004 = 1$, donc si $\widetilde{X} \hookrightarrow \mathcal{P}(1)$, on a

$$\begin{aligned}
\mathbb{P}(X \leq 2) &= \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) \\
&\simeq \mathbb{P}(\widetilde{X}=0) + \mathbb{P}(\widetilde{X}=1) + \mathbb{P}(\widetilde{X}=2) \\
&= e^{-1} + e^{-1} + \frac{e^{-1}}{2} \\
&= \frac{5}{2}e^{-1}.
\end{aligned}$$

Remarque 4.5. Dans la pratique, ce résultat nous permet de remplacer la loi $\mathcal{B}(n,p)$ par la loi $\mathcal{P}(np)$ lorsque le produit np est très petit. L'avantage de la loi de Poisson est qu'elle ne dépend que d'un seul paramètre. Bien évidement ceci reste une approximation, mais elle est acceptable dès que $n > 50$ et $p < 0.1$.

4.4 Transformation d'une variable aléatoire

Commençons par introduire la notion de support pour une variable aléatoire

Définition 4.14. Soit X une variable aléatoire réelle.

- Si X est discrète de fonction de masse p_X, alors on appelle support de X l'ensemble $\{x \in \mathbb{R},\ p_X(x) \neq 0\}$.
- Si X est absolument continue de fonction de densité f_X, alors on appelle support de X l'ensemble $\{x \in \mathbb{R},\ f_X(x) \neq 0\}$.

Le support d'une variable aléatoire X est noté $supp(X)$

Dans plusieurs situations, on est amené à utilisér non pas une variable aléatoire X de loi usuelle mais une fonction de celle-ci, i. e. $Y = h(X)$ où h est une fonction réelle. il devient donc naturel de chercher une méthode qui permet de déduire les principales caractéristique la nouvelle variable Y à partir de celles de X. Dans cette section on propose une méthode pour le faire à partir de 4 exemples, en distingant le cas bijectif et la cas non-bijectif.

Tout au long de ce paragraphe $h:\ \mathbb{R} \to \mathbb{R}$ est une fonction.

4.4.1 Le cas discret

Théorème 4.2. Soient X une variable aléatoire discrète, p_X sa fonction de masse et $supp(X)$ son support. On pose $Y = h(X)$, alors Y est une variable aléatoire discrète, son support $supp(Y) = h\Big(supp(X)\Big)$ et sa fonction de masse p_Y est donnée par

$$p_Y(y) = \begin{cases} \sum_{\{k/\ h(k)=y\}} p_X(k) & \text{si } y \in supp(Y) \\ 0 & \text{sinon} \end{cases}$$

Exemple 4.4. Soit X une variable aléatoire discrète qui prend les valeurs $-1, 0, 1, 2$ avec les probabilités respectives $\frac{1}{8},\ \frac{1}{16},\ \frac{1}{4},\ \frac{9}{16}$. Posons $h(x) = |x|$, on a dans ce cas $supp(X) = \{-1, 0, 1, 2\}$ et $supp(Y) = \{0, 1, 2\}$. Par suite :

$$p_Y(y) = \begin{cases} \frac{1}{16} & \text{si } y = 0 \\ \frac{3}{8} & \text{si } y = 1 \\ \frac{9}{16} & \text{si } y = 2 \\ 0 & \text{sinon} \end{cases}$$

4.4.2 Le cas absolument continu

Théorème 4.3. Soient X une variable aléatoire absolument continue, f_X sa fonction de densité et $supp(X)$ son support. On pose $Y = h(X)$ et on suppose que la fonction h soit bijective sur $supp(X)$, et que sa fonction réciproque h^{-1} admette une dérivée continue sur $h(supp(X))$. Alors Y est une variable aléatoire absolument continue, son support $supp(Y) = h\Big(supp(X)\Big)$ et sa fonction de densité f_Y est donnée par

$$f_Y(y) = \begin{cases} f_X(h^{-1}(y))|(h^{-1})'(y)| & \text{si } y \in supp(Y) \\ 0 & \text{sinon} \end{cases}$$

Exemple 4.5. Soient $X \hookrightarrow \mathcal{U}[1, 3]$, $h(x) = e^x$ et $Y = h(X)$. Il est clair que $supp(X) = [1, 3], supp(Y) = [e, e^3]$ et h est bijective sur $[1, 3]$, puisque $h^{-1}(x) = \ln(x)$, de plus h^{-1} admet une dérivée continue sur $[1, 3]$, ainsi la fonction de densité f_Y est donnée par

$$f_Y(y) = f_X(\ln(y))\frac{1}{y} = \frac{1}{2y}\mathbb{1}_{[e,e^3]}(y).$$

Exemple 4.6. Soient $\lambda > 0$, $X \hookrightarrow \mathcal{E}(\lambda)$, $h(x) = \frac{1}{x}$ et $Y = h(X)$. On peut voir que $supp(X) = [0, +\infty[$, $supp(Y) =]0, +\infty[$ et h est bijective sur $]0, +\infty[$, puisque $h^{-1}(x) = \frac{1}{x}$, de plus h^{-1} admet une dérivée continue sur $]0, +\infty[$, ainsi la fonction de densité f_Y est donnée par

$$f_Y(y) = f_X(\frac{1}{y})|\frac{-1}{y^2}| = \frac{\lambda}{y^2}e^{-\frac{\lambda}{y}}\mathbb{1}_{]0,+\infty[}(y).$$

Remarque 4.6. L'hypothèse de bijectivité de la fonction h sur tout le support de X est trop forte, en effet même si cette hypothèse n'est pas vérifiée on peut dans la plupart des cas partager le support de X en plusieurs parties de telle sorte que h soit bijective dans chaque partie, ainsi on peut appliquer le théorème sur chaque partie, pour le voir considérons :

Exemple 4.7. Soient $X \hookrightarrow \mathcal{N}(0,1)$, $h(x) = x^2$ et $Y = h(X)$. Dans ce cas $supp(X) = \mathbb{R}$, ainsi h n'est pas bijective sur $supp(X)$. Soit F_Y la fonction de répartition de Y et remarquons que $supp(Y) = [0, +\infty[$. Soit $y \in supp(Y)$, on peut écrire

$$\begin{aligned} F_Y(y) &= \mathbb{P}\Big(Y \leq y\Big) \\ &= \mathbb{P}\Big(X^2 \leq y\Big) \\ &= \mathbb{P}\Big(-\sqrt{y} \leq X \leq \sqrt{y}\Big) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

D'où

$$f_Y(y) = F'_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + f_X(\sqrt{y}) \frac{1}{2\sqrt{y}},$$

par conséquent

$$f_Y(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{1}{2}y} \, \mathbb{1}_{[0,+\infty[}(y).$$

4.5 Exercices

Exercice 4.1. Soit X une variable aléatoire réelle, trouver $\mathbb{E}(X)$, $Var(X)$ et σ_X dans les cas suivants :

- $X \hookrightarrow \mathcal{B}(p)$.
- $X \hookrightarrow \mathcal{B}(n, p)$.
- $X \hookrightarrow \mathcal{P}(\lambda)$.
- $X \hookrightarrow \mathcal{G}(p)$.
- $X \hookrightarrow \mathcal{U}_{[a,b]}$.
- $X \hookrightarrow \mathcal{E}(\lambda)$.
- $X \hookrightarrow \Gamma(\lambda, r)$.
- $X \hookrightarrow \mathcal{L}$.
- $X \hookrightarrow \mathcal{N}(0, 1)$.
- $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

Exercice 4.2. Dans un examen de type QCM, on pose 10 questions. Pour chaque question on propose au candidat 3 réponses dont une seule est juste. Une personne ignorant totalement le sujet se présente à cet examen et coche les réponses pour chaque question au hasard. Soit X la variable aléatoire qui donne le nombre de réponses justes.

- Quelle est la loi de X ?
- Quelle est la probabilité que cette personne donne 3 réponses justes ?
- Une personne est admise si elle donne au moins 5 réponses justes. Quelle est la probabilité que cette personne soit admise ?

Exercice 4.3. Les clients d'un magasin arrivent selon une loi de Poisson avec une moyenne de deux clients par heure. Qelle est la loi de la variable X qui donne le temps nécessaire pour que dix clients arrivent ?

Exercice 4.4. La durée de vie d'un composant éléctrique suit une loi exponentielle de moyenne 5 ans. Sachant que ce composant a fonctionné pendant une année, quelle est la probabilité qu'il tombe en panne dans les quatres années suivantes.

Exercice 4.5. Déterminer la fonction de répartition F_X pour la variable aléatoire X dans les cas suivants :

- $X \hookrightarrow \mathcal{B}(p)$.
- $X \hookrightarrow \mathcal{U}_{[a,b]}$.
- $X \hookrightarrow \mathcal{E}(\lambda)$.
- $X \hookrightarrow \mathcal{L}$.

Exercice 4.6. Soit F la fonction de répartition d'une certaine variables aléatoire X. Supposons que F est continue et posons $Y = F(X)$. Montrer que Y suit une loi uniforme sur $[0, 1]$.

Exercice 4.7. Soient $\lambda > 0$, $\{p_n,\ n \geq 1\}$ une suite de nombres réels vérifiant : :

$$\forall n \geq 1,\ 0 < p_n < 1, \qquad \text{et} \qquad \lim_{n \to +\infty} np_n = \lambda,$$

Montrer que pour tout $z \in \mathbb{C}$:

$$\lim_{n \to +\infty} \Big(1 - p_n z\Big)^n = e^{\lambda z}.$$

Exercice 4.8. Pour $r > 0$, on définit la fonction Γ par

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

- Caluler $\Gamma(0)$ et $\Gamma(1)$.
- Montrer que pour tout $r > 0$, $\Gamma(r+1) = r\Gamma(r)$.
- En déduire que pour tout $n \geq 1$, $\Gamma(n) = (n-1)!$.

Exercice 4.9. Soient $\lambda > 0, t_1, t_2 \in \mathbb{R}$ et X une variable aléatoire de loi $\mathcal{E}(\lambda)$. Montrer que

$$I\!P(X > t_1 + t_2 / X > t_1) = I\!P(X > t_2).$$

Exercice 4.10. Soient $\mu \in \mathbb{R}$, $\sigma > 0$ et $X \hookrightarrow \mathcal{N}(0, 1)$.

1. Posons $Y = \sigma X + \mu$, calculer F_Y la fonction de répartition de Y en fonction de F_X et en déduire la densité de Y. Identifier la loi de Y.

2. Soit à présent U une variable aléatoire de loi $\mathcal{N}(\mu, \sigma^2)$, posons $V = \frac{U-\mu}{\sigma}$. Calculer F_V la fonction de répartition de V en fonction de F_U et en déduire la densité de V. Identifier la loi de V.

Exercice 4.11. Deux espèces de poissons vivent dans un petit barrage, 200 de la première espèce et 60 de la seconde. On pêche 5 poissons de ce barrage. Calculer la probabilité de ne pas pêcher de poissons de la seconde espèce.

Exercice 4.12. Les résultats d'un test de quotient intellectuel pour une classe dans un collège montre que le QI de cette classe suit une loi normale de moyenne 100 et de variance 225. Quel est le pourcentage des élève ayant un QI inférieur à 91 ou supérieur à 130.

Exercice 4.13. Dans un village, la taille des hommes peut être modélisée par une variable aléatoire normale de moyenne 167 cm et d'écart-type $3cm$. Calculer le pourcentage des hommes qui ont une taille

– supérieur à 170 cm.
– inférieur à 165 cm
– comprise entre 162 cm et 172 cm.

On choisit au hasard 4 hommes de ce village, quelle est la probabilité que

– Les 4 hommes aient une taille supérieur à 170 cm.
– Deux aient une taille supérieure à la moyenne et deux ont une taille inférieure à la moyenne.

Exercice 4.14. Soit X une variable aléatoire réelle absolument continue de fonction de densité

$$f_X(x) = \frac{1}{2}e^{-|x-1|}, \qquad x \in \mathbb{R}.$$

1. Calculer la fonction de répartition, la moyenne et la variance de X.
2. On pose $Y = 2X + 1$.

i. Calculer F_Y, la fonction de répartition de Y et en déduire f_Y la fonction de densité de Y.

ii. Calculer $\mathbb{E}(Y)$ et $Var(Y)$ par deux méthodes différentes.

Exercice 4.15. Soient $\lambda > 0,\ X \hookrightarrow \mathcal{P}(\lambda)$.

1. Vérifier que la v. a. r $\frac{1}{1+X}$ est intégrable.
2. Calculer $\mathbb{E}(\frac{1}{1+X})$ et $\mathbb{E}(\frac{1+X}{2+X})$.
3. En déduire $\mathbb{E}(\frac{1}{2+X})$.

Chapitre 5

Couples aléatoires

Au chapitre 3 nous avons introduit la notion de variable aléatoire à une dimension, le but de ce chapitre est de généraliser cette notion aux variables à deux dimensions (couple). Toutes les variables aléatoires de ce chapitre sont définies sur un espace de probabilité $(\Omega, \mathcal{F}, I\!P)$.

5.1 Définitions-Exemples

Définition 5.1. Soient X, Y deux applications de Ω vers $\mathbb{R}$. Si X et Y sont deux variables aléatoires réelles alors le couple (X, Y) est appelé couple aléatoire.

Exemple 5.1. On jette deux dés équilibrés et on note par X le chiffre obtenu par le premier dé, Y la somme des chiffres obtenus. (X, Y) est un couple aléatoire.

Exemple 5.2. On prélève un groupe de TD au hasard et on mesure le poids et la taille des étudiants de ce groupe. On note par X la taille en cm et Y le poids en kg. (X, Y) est un couple aléatoire.

Définition 5.2. Soit (X, Y) un couple aléatoire, l'application $F_{(X,Y)}$ définie sur $\mathbb{R} \times \mathbb{R}$ par

$$\forall x, y \in \mathbb{R}, \qquad F_{(X,Y)}(x, y) = I\!P\Big(X \leq x \ , \ Y \leq y\Big).$$

est applée la <u>fonction de répartition</u> du couple (X, Y).

Comme dans le cas réel, la fonction de répartition possède de nombreuses bonnes propriétés :

Proposition 5.1. Soient (X,Y) un couple aléatoire, $F_{(X,Y)}$ sa fonction de répartition, alors

- Pour tout $(x,y) \in \mathbb{R} \times \mathbb{R}, \quad 0 \leq F_{(X,Y)}(x,y) \leq 1$.
- $\lim_{x \text{ et } y \to -\infty} F_{(X,Y)}(x,y) = 0$ et $\lim_{x \text{ et } y \to +\infty} F_{(X,Y)}(x,y) = 1$.
- Pour tout $y \in \mathbb{R}$, l'application $x \longrightarrow F_{(X,Y)}(x,y)$ est croissante et continue à droite.
- Pour tout $x \in \mathbb{R}$, l'application $y \longrightarrow F_{(X,Y)}(x,y)$ est croissante et continue à droite.

Définition 5.3. Soit (X,Y) un couple aléatoire, Les fonctions de répartition marginales associées au couple (X,Y) sont données par :

$$F_X(x) = \lim_{y \to +\infty} F_{(X,Y)}(x,y)$$

et

$$F_Y(y) = \lim_{x \to +\infty} F_{(X,Y)}(x,y)$$

5.2 Couples discrets

Définition 5.4. Soient X, Y deux applications de Ω vers $\mathbb{R}$. On dit que (X,Y) est un couple aléatoire discret si X et Y sont deux variables aléatoire discrètes.

Définition 5.5. Soit (X,Y) un couple aléatoire discret, l'application $p_{(X,Y)}$ définie sur $\mathbb{R} \times \mathbb{R}$ par

$$p_{(X,Y)}(x,y) = \begin{cases} \mathbb{P}(X = x, Y = y) & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

est applée la fonction de masse du couple (X,Y).

Exemple 5.3. On jette deux dés équilibrés et on note par X le chiffre obtenu par le premier dé, Y la variable aléatoire qui vaut 1 si le deuxième dé ramène un chiffre inférieur ou égale à 4 et 0 sinon. La fonction de masse pour ce couple aléatoire est donnée par

$$p_{(X,Y)}(x,y) = \begin{cases} \frac{2}{3} \text{ si } (x,y) \in \{1,2,3,4\} \times \{1\} \\ \frac{1}{3} \quad \text{si } (x,y) \in \{5,6\} \times \{0\} \\ 0 \quad \text{sinon} \end{cases}$$

Proposition 5.2. Une application p définie sur $\mathbb{R} \times \mathbb{R}$ est une fonction de masse si et seulement si

– Pour tout $(x,y) \in \mathbb{R} \times \mathbb{R}, \quad p(x,y) \geq 0$.
– $\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} p(x,y) = 1$.

Définition 5.6. Soit (X,Y) un couple aléatoire discret, $p_{(X,Y)}$ sa fonction de masse et $supp(X)$ (resp. $supp(Y)$) le support de X (resp. le support de Y).

– <u>Les fonctions de masse marginales</u> associées au couple (X,Y) sont données par :

$$p_X(x) = \begin{cases} \sum_{y \in supp(Y)} p_{X,Y}(x,y) & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

et

$$p_Y(y) = \begin{cases} \sum_{x \in supp(X)} p_{X,Y}(x,y) & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

Remarque 5.1. Les fonctions de masse marginales sont aussi appelées les lois marginales associées au couple (X,Y).

Définition 5.7. Soient (X,Y) un couple aléatoire discret, $p_{(X,Y)}$ sa fonction de masse. <u>l'espérance mathématique</u> ou <u>la moyenne</u> du couple (X,Y) est $(\mathbb{E}(X), \mathbb{E}(Y))$.

Définition 5.8. Soient (X,Y) un couple aléatoire discret, $p_{(X,Y)}$ sa fonction de masse.

– la covariance du couple (X,Y) est donnée par

$$Cov(X,Y) = \sum_{(i,j)\in supp(X)\times supp(Y)} ij p_{(X,Y)}(i,j) \; - \; \mathbb{E}(X)\mathbb{E}(Y).$$

– Supposons que $Var(X) > 0$ et $Var(Y) > 0$. Le coefficient de corrélation du couple (X,Y) est donnée par

$$\rho_{(X,Y)} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

Proposition 5.3. Soit (X,Y) un couple aléatoire discret, alors

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

Définition 5.9. Soient (X,Y) un couple aléatoire discret, $p_{(X,Y)}$ sa fonction de masse et $supp(X)$ (resp. $supp(Y)$) le support de X (resp. le support de Y).

– La loi conditionnelle de X sachant Y est donnée par :

$$p_{X/Y=y}(x) = \begin{cases} \frac{p_{(X,Y)}(x,y)}{p_Y(y)} & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

– La loi conditionnelle de Y sachant X est donnée par :

$$p_{X/Y=y}(x) = \begin{cases} \frac{p_{(X,Y)}(x,y)}{p_X(x)} & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

Définition 5.10. Soit (X,Y) un couple aléatoire discret.

– L'espérance conditionnelle de X sachant Y est donnée par :

$$\forall y \in supp(Y), \qquad \mathbb{E}\Big(X/Y=y\Big) = \sum_{i\in supp(X)} i p_{X/Y=y}(i).$$

– L'espérance conditionnelle de Y sachant X est donnée par :

$$\forall x \in supp(X), \qquad \mathbb{E}\Big(Y/X=x\Big) = \sum_{j\in supp(Y)} j p_{Y/X=x}(j).$$

5.3 Couples absolument continus

Définition 5.11. Un couple aléatoire (X, Y) est absolument continu si sa fonction de répartition $F_{(X,Y)}$ est continue à droite et à gauche par rapport à chacune des variables et possède une dérivée seconde :

$$\frac{\partial^2 F_{(X,Y)}}{\partial x \partial y}$$

sur un domaine non-vide $\mathcal{D}$.

Définition 5.12. Soit (X, Y) un couple aléatoire absolument continu. La fonction

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{\partial^2 F_{(X,Y)}}{\partial x \partial y}(x, y) & \text{si } (x, y) \in \mathcal{D} \\ 0 & \text{sinon} \end{cases}$$

est appelée la <u>fonction de densité</u> du couple (X, Y).

Proposition 5.4. Une application f définie sur $\mathbb{R} \times \mathbb{R}$ est une fonction de densité si et seulement si

- Pour tout $(x, y) \in \mathbb{R} \times \mathbb{R}, \quad f(x, y) \geq 0$.
- $\int_{\mathbb{R}^2} f(x, y) dx dy = 1$.

Définition 5.13. Soit (X, Y) un couple aléatoire absolument continu, $f_{(X,Y)}$ sa fonction de densité et *supp*(X) (resp. *supp*(Y)) le support de X (resp. le support de Y).

- <u>Les fonctions de densité marginales</u> associées au couple (X, Y) sont données par :

$$f_X(x) = \begin{cases} \int_{\mathbb{R}} f_{X,Y}(x, y) dy & \text{si } (x, y) \in \mathcal{D} \\ 0 & \text{sinon} \end{cases}$$

et

$$f_Y(y) = \begin{cases} \int_{\mathbb{R}} f_{X,Y}(x, y) dx & \text{si } (x, y) \in \mathcal{D} \\ 0 & \text{sinon} \end{cases}$$

– Les fonctions de répartition marginales associées au couple (X, Y) sont données par :

$$F_X(x) = \lim_{y\to+\infty} F_{(X,Y)}(x, y)$$

et

$$F_Y(y) = \lim_{x\to+\infty} F_{(X,Y)}(x, y)$$

Remarque 5.2. Les fonctions de densité marginales sont aussi appelées les lois marginales associées au couple (X, Y).

Proposition 5.5. Si (X, Y) est un couple aléatoire absolument continu, alors pour tout $(x, y) \in \mathcal{D}$:

$$f_X(x) = F_X^{'}(x), \qquad \text{et} \qquad f_Y(y) = F_Y^{'}(y).$$

Exemple 5.4. Soient $c \in \mathbb{R}, f$ la fonction définie par

$$f(x, y) = cxye^{-x^2-y^2}, \qquad x,\ y \in \mathbb{R}^+.$$

– Déterminer c pour que f soit la fonction de densité d'un couple aléatoire (X, Y).
– Calculer $F_{(X,Y)}$ la fonction de répartition du couple (X, Y).
– En déduire les fonctions de répartitions marginales.
– Calculer les fonctions de densité marginale.

SOLUTION:

– Il faut remarquer d'abord que f doit être positive, ainsi $c \geq 0$, de plus

$$1 = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x, y)dxdy = c\int_R xe^{-x^2}dx\int_R ye^{-y^2}dy = \frac{c}{4}.$$

Ainsi $c = 4$.

– Soient $x,\ y \in \mathbb{R}^+$, la fonction de répartition du couple est donnée par

$$\begin{aligned} F_{(X,Y)}(x,y) &= \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv \\ &= 4\int_{-\infty}^{x} ue^{-u^2}du \int_{-\infty}^{y} ve^{-v^2}dv \\ &= \left(1-e^{-x^2}\right)\left(1-e^{-y^2}\right). \end{aligned}$$

D'autre part $F_{(X,Y)}$ est nulle lorsque l'une des coordonnées x ou y est négative.

– Pour les fonctions de répartitions marginales, il suffit de prendre la limite, en effet

$$F_X(x) = \lim_{y\to+\infty}\left(1-e^{-x^2}\right)\left(1-e^{-y^2}\right)\mathbb{1}_{\mathbb{R}^+\times\mathbb{R}^+}(x,y) = \left(1-e^{-x^2}\right)\mathbb{1}_{\mathbb{R}^+}(x)$$

et

$$F_Y(y) = \lim_{x\to+\infty}\left(1-e^{-x^2}\right)\left(1-e^{-y^2}\right)\mathbb{1}_{\mathbb{R}^+\times\mathbb{R}^+}(x,y) = \left(1-e^{-y^2}\right)\mathbb{1}_{\mathbb{R}^+}(y).$$

– Pour les fonctions de densité marginales, on a

$$f_X(x) = F_X^{'}(x) = 2xe^{-x^2}\mathbb{1}_{R^+}(x).$$

et

$$f_Y(y) = F_Y^{'}(y) = 2ye^{-y^2}\mathbb{1}_{R^+}(y).$$

Définition 5.14. Soient (X,Y) un couple aléatoire absolument continu, $f_{(X,Y)}$ sa fonction de densité. l'espérance mathématique ou la moyenne du couple (X,Y) est $(\mathbb{E}(X),\mathbb{E}(Y))$.

Définition 5.15. Soient (X,Y) un couple aléatoire absolument continu, $f_{(X,Y)}$ sa fonction de densité.

– La covariance du couple (X,Y) est donnée par

$$Cov(X,Y) = \int_{\mathbb{R}^2} xyf_{(X,Y)}(x,y)dxdy \ - \ \mathbb{E}(X)\mathbb{E}(Y).$$

– Supposons que $Var(X) > 0$ et $Var(Y) > 0$. Le coefficient de corrélation du couple (X,Y) est donnée par

$$\rho_{(X,Y)} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

Proposition 5.6. Soit (X,Y) un couple aléatoire absolument continu, alors

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y).$$

Définition 5.16. Soient (X,Y) un couple aléatoire absolument continu, $f_{(X,Y)}$ sa fonction de densité et *supp*(X) (resp. *supp*(Y)) le support de X (resp. le support de Y).

– La loi conditionnelle de X sachant Y est donnée par :

$$f_{X/Y=y}(x) = \begin{cases} \frac{f_{(X,Y)}(x,y)}{f_Y(y)} & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

– La loi conditionnelle de Y sachant X est donnée par :

$$f_{X/Y=y}(x) = \begin{cases} \frac{f_{(X,Y)}(x,y)}{f_X(x)} & \text{si } (x,y) \in supp(X) \times supp(Y) \\ 0 & \text{sinon} \end{cases}$$

Définition 5.17. Soit (X,Y) un couple aléatoire absolument continu.

– L'espérance conditionnelle de X sachant Y est donnée par :

$$\forall y \in supp(Y), \qquad \mathbb{E}\Big(X/Y=y\Big) = \int_{supp(X)} x f_{X/Y=y}(x)dx.$$

– L'espérance conditionnelle de Y sachant X est donnée par :

$$\forall x \in supp(X), \qquad \mathbb{E}\Big(Y/X=x\Big) = \int_{supp(Y)} y f_{Y/X=x}(y)dy.$$

Exemple 5.5. Soit (X,Y) un couple aléatoire de fonction de densité

$$f(x,y) = \frac{3}{2}\Big(x^2+y^2\Big)\mathbb{1}_{[0,1]\times[0,1]}(x,y).$$

– Calculer les densités marginales.

– Déterminer les lois conditionnelles.

SOLUTION:

– Pour les densité marginales, on a pour $x, y \in [0,1]$:

$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy = \frac{3}{2}\Big(x^2 + \int_0^1 y^2 dy\Big) = \frac{3}{2}x^2 + \frac{1}{2}.$$

De même

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)dx = \frac{3}{2}\Big(y^2 + \int_0^1 x^2 dy\Big) = \frac{3}{2}y^2 + \frac{1}{2}.$$

Ainsi

$$f_X(x) = \frac{3}{2}x^2 + \frac{1}{2}\mathbb{1}_{[0,1]}(x) \qquad \text{et} \qquad f_Y(y) = \frac{3}{2}y^2 + \frac{1}{2}\mathbb{1}_{[0,1]}(y).$$

– Pour les lois conditionnelles, soient $x, y \in [0,1]$, on a

$$f_{X/Y=y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{3(x^2+y^2)}{3y^2+1}.$$

De même

$$f_{Y/X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{3(x^2+y^2)}{3x^2+1}.$$

Exemple 5.6. Soit (X,Y) un couple aléatoire de fonction de densité

$$f_{(X,Y)}(x,y) = e^{-y}\mathbb{1}_{\mathbb{R}_+ \times\ [x,+\infty[}(x,y).$$

– Donner les lois marginales des v. a X, Y.
– Déterminer $f_{X/Y=y}$ et $f_{Y/X=x}$.
– En déduire $E(X/Y)$ et $E(Y/X)$.

SOLUTION:

– Commençons par la loi marginale selon X, soit $x \in \mathbb{R}_+$, on a

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)dy = \int_x^{+\infty} e^{-y}dy = e^{-x},$$

d'où

$$f_X(x) = e^{-x}\mathbb{1}_{[0,+\infty[}(x).$$

– De même, si $y \in \mathbb{R}_+$, on a

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{(X,Y)}(x,y)dx = \int_0^y e^{-y}dx = ye^{-y},$$

d'où

$$f_Y(y) = ye^{-y}\mathbb{1}_{[0,+\infty[}(y).$$

– Calculons à présent les lois conditionnelles, soient $x \in \mathbb{R}_+, 0 < y \leq x$:

$$f_{X/Y=y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y},$$

De mmême

$$f_{Y/X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}.$$

Par suite

$$f_{X/Y=y}(x) = \frac{1}{y}\mathbb{1}_{]0,y]}(x) \qquad \text{et} \qquad f_{Y/X=x}(y) = e^{-(y-x)}\mathbb{1}_{[x,+\infty[}(y).$$

– Pour les espérances conditionnelles, on a pour $y \in \mathbb{R}_+$:

$$\mathbb{E}(X/Y=y) = \int_{-\infty}^{+\infty} xf_{X/Y=y}(x)dx = \int_0^y \frac{x}{y}dx = \frac{y}{2},$$

De même si $x \in \mathbb{R}_+$, on faisant une intégration par parties, on obtient

$$\mathbb{E}(Y/X=x) = \int_{-\infty}^{+\infty} yf_{Y/X=x}(y)dx = e^x\int_x^{+\infty} ye^{-y}dy = e^x\Big(xe^{-x}+e^{-x}\Big) = x+1.$$

Ainsi

$$\mathbb{E}(X/Y) = \frac{Y}{2} \qquad \text{et} \qquad \mathbb{E}(Y/X) = X+1.$$

5.4 Variables aléatoires indépendantes

La notion d'indépendance est très importante en théorie des probabilités, elle est l'hypothèse principale dans plusieurs résultats fondamentaux concernant le comportement asymptotique des moyennes de variables aléatoires. Commençons par rappeler la notion de tribu engendrée par une variable aléatoire définie dans le troisième chapitre. Soit X une variable aléatoire réelle, la tribu engendrée par X et notée $\sigma(X)$ est définie par

$$\sigma(X) = \{X^{-1}(A),\ A \in \mathcal{B}_{\mathbb{R}}\},$$

où $\mathcal{B}_{\mathbb{R}}$ est la tribu boréliènne sur $\mathbb{R}$.

Définition 5.18. Soient $X,\ Y$ deux variables aléatoires réelles. On dit que X et Y sont indépendantes si les tribus engendrées par les variables X et Y sont indépendantes. i. e.

$$\forall A,\ B \in \mathcal{B}_{\mathbb{R}}, \qquad I\!\!P\Big(X \in A, Y \in B\Big) = I\!\!P\Big(X \in A\Big)I\!\!P\Big(Y \in B\Big).$$

Remarque 5.3.
- Vérifier l'indépendance de deux variables aléatoires revient à vérifier l'indépendance des tribus engendrées par ces deux variables.
- De la même manière, la notion d'indépendance ainsi définie peut être généralisée à plusieurs variables.

En général, il est difficile de prouver l'indépendance de deux variables aléatoires en utilisant cette définition, c'est pour cette raison que nous allons proposer d'autre critères.

Proposition 5.7. Soient (X, Y) un couple aléatoire, $F_{(X,Y)}$ sa fonction de répartition, F_X et F_Y les fonctions de répartition marginales. Alors

$$\text{Les variables } X \text{ et } Y \text{ sont indépendantes} \Longleftrightarrow F_{(X,Y)}(x, y) = F_X(x)F_Y(y).$$

Dans le cas où on connait la nature du couple (X, Y) (*discret ou absolument continu*), on a

Proposition 5.8. Soit (X,Y) un couple aléatoire, vérifiant :

$$Supp(X,Y) = Supp(X) \times Supp(Y).$$

– Si (X,Y) est un couple discret de fonction de masse $p_{(X,Y)}$, alors les variables X et Y sont indépendantes si et seulement si

$$\forall (x,y) \in supp(X) \times supp(Y), \qquad p_{(X,Y)}(x,y) = p_X(x) p_Y(y),$$

– Si (X,Y) est un couple absolument continu de fonction de densité $f_{(X,Y)}$, alors les variables X et Y sont indépendantes si et seulement si

$$\forall (x,y) \in supp(X) \times supp(Y), \qquad f_{(X,Y)}(x,y) = f_X(x) f_Y(y),$$

Exemple 5.7. Soient $\lambda,\ \mu \in \mathbb{R}^+,\ (X,Y)$ un couple aléatoire de fonction de densité

$$f(x,y) = \lambda\mu e^{-(\lambda x + \mu y)} \mathbb{1}_{\mathbb{R}^+ \times \mathbb{R}^+}(x,y).$$

On remarque que $Supp(X,Y) = \mathbb{R}^+ \times \mathbb{R}^+ = Supp(X) \times Supp(Y)$, de plus on peut écrire

$$f(x,y) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x) \times \mu e^{-\mu y} \mathbb{1}_{\mathbb{R}^+}(y).$$

Par suite, les variables aléatoires X et Y sont indépendantes.

L'indépendance peut être vérifiée aussi en utilisant la fonction génératrice des moments et la fonction caractéristique, commençons d'abord par définir ces fonction pour les couples aléatoires.

Définition 5.19. Soit (X,Y) un couple aléatoire.

– La fonction génératrice des moments associée au couple (X,Y) est définie par

$$\forall s,t \in \mathbb{R}, \qquad M_{(X,Y)}(s,t) = \mathbb{E}\Big(e^{sX+tY}\Big).$$

– La fonction caractéristique associée au couple (X,Y) est définie par

$$\forall s,t \in \mathbb{R}, \qquad \varphi_{(X,Y)}(s,t) = \mathbb{E}\Big(e^{isX+itY}\Big),$$

où i est le nombre complexe vérifiant : $i^2 = -1$.

Proposition 5.9. Soient (X, Y) est un couple aléatoire, $M_{(X,Y)}$ sa foncton génératrice des moments et $\varphi_{(X,Y)}$ sa fonction caractéristique. Les trois propriétés suivantes sont équivalentes :

- X et Y sont indépendantes.
- Pour tout $(s,t) \in \mathbb{R}^2,\ M_{(X,Y)}(s,t) = M_X(s)M_Y(t)$.
- Pour tout $(s,t) \in \mathbb{R}^2,\ \varphi_{(X,Y)}(s,t) = \varphi_X(s)\varphi_Y(t)$.

Une conséquence importante de l'indépendance de deux variables aléatoires est le résultat suivant :

Proposition 5.10. Soient X, Y deux variables aléatoires indépendantes, alors

- Pour tout couple de fonctions continues (g, h), les variables aléatoires $g(X)$ et $h(Y)$ sont aussi indépendantes.
- Pour tout couple de fonctions continues (g, h),

$$\mathbb{E}\Big(g(X)h(Y)\Big) = \mathbb{E}\Big(g(X)\Big)\mathbb{E}\Big(h(Y)\Big).$$

  En particulier $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.
- $Cov(X, Y) = 0$.
- $Var(X + Y) = Var(X) + Var(Y)$.
- Pour tout $t \in \mathbb{R},\quad M_{X+Y}(t) = M_X(t)M_Y(t)$.
- Pour tout $t \in \mathbb{R},\quad \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

Exemple 5.8. On peut se servir de ce dernier résultat pour calculer la loi de la somme de deux variables aléatoires indépendantes, en effet considérons les deux exemples suivants :

- Soient $n, m \geq 1,\ 0 < p < 1$ et X, Y deux variables aléatoires indépendantes avec $X \hookrightarrow \mathcal{B}(n,p),\ Y \hookrightarrow \mathcal{B}(m,p)$. Posons $S = X+Y$ et soient $t \in \mathbb{R},\ M_S$ la fonction génératrice des moments de S. On a par cette dernière proposition et la Proposition 4.5 du Chapitre 4 :

$$M_S(t) = M_{X+Y}(t) = M_X(t)M_Y(t) = \Big(1 - p + pe^t\Big)^n\Big(1 - p + pe^t\Big)^m.$$

On a donc, pour tout $t \in \mathbb{R}$:

$$M_S(t) = \left(1 - p + pe^t\right)^{n+m}.$$

En faisant appel encore une fois à la Proposition 4.5 du quatrième chapitre, on conclut que

$$S \hookrightarrow \mathcal{B}(n+m, p).$$

– Soient $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 \in \mathbb{R}_+^*$ et X, Y deux variables aléatoires indépendantes avec $X \hookrightarrow \mathcal{N}(\mu, \sigma_1^2)$, $Y \hookrightarrow \mathcal{N}(\mu_2, \sigma_2^2)$. Posons $S = X + Y$ et soient $t \in \mathbb{R}$, φ_S la fonction caractéristique de S. On a par cette dernière proposition et la Proposition 4.24 du Chapitre 4 :

$$\varphi_S(t) = \varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = e^{i\mu_1 t - \frac{\sigma_1^2}{2}t^2} e^{i\mu_2 t - \frac{\sigma_1^2}{2}t^2}.$$

On obtient alors

$$\varphi_S(t) = e^{i(\mu_1+\mu_2)t - \frac{(\sigma_1^2+\sigma_2^2)}{2}t^2},$$

pour tout $t \in \mathbb{R}$. Par la Proposition 4.24 du quatrième chapitre, on peut affirmer que

$$S \hookrightarrow \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

Un autre avantage de l'indépendance de deux variables aléatoires X et Y, est qu'elle permet dans plusieurs situations de calculer explicitement la loi de la somme $X + Y$, lorsque on connait les lois des variables X et Y, en effet

Proposition 5.11. Soient X et Y deux variables aléatoires indépendantes, posons $S = X + Y$.

– Si X et Y sont deux variables discrètes de fonctions de masse p_X et p_Y, alors S est aussi une variable aléatoire discrète de fonction de masse p_S donnée par

$$p_S(k) = \sum_{(k-j,j)\in Supp(X)\times Supp(Y)} p_X(k-j)p_Y(j)$$

– Si X et Y sont deux variables absolument continues, de fonctions de densité

f_X et f_Y, alors S est aussi une variable aléatoire absolument coninue de fonction de densité f_S donnée par

$$f_S(x) = \int_{-\infty}^{+\infty} f_X(x-y)f_Y(y)dy.$$

Exemple 5.9. Soient $\lambda > 0, \mu > 0,\ X, Y$ deux variables aléatoires indépendantes avec $X \hookrightarrow \mathcal{P}(\lambda),\ Y \hookrightarrow \mathcal{P}(\mu)$. On pose $S = X + Y$. Trouver la loi de S.

<u>SOLUTION</u>:

D'après la proposition précédente, S est une variable aléatoire discrète, de plus $Supp(S) = \mathbb{N}$. Ainsi pour $k \in \mathbb{N}$ on a

$$\begin{aligned} p_S(k) &= \sum_{(k-j,j)\in Supp(X)\times Supp(Y)} p_X(k-j)p_Y(j) \\ &= e^{-(\lambda+\mu)} \sum_{(k-j,j)\in\mathbb{N}\times\mathbb{N}} \frac{\lambda^{k-j}\mu^j}{(k-j)!j!} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^{n} \frac{k!}{(k-j)!j!}\lambda^{k-j}\mu^j \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{j=0}^{n} C_k^j \lambda^{k-j}\mu^j \\ &= e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^k}{k!}. \end{aligned}$$

Par conséquent $S \hookrightarrow \mathcal{P}(\lambda + \mu)$.

Exemple 5.10. Soient X, Y deux variables aléatoires indépendantes avec $X \hookrightarrow \mathcal{E}(1)$ et $Y \hookrightarrow \mathcal{E}(1)$. On pose $S = X + Y$. Trouver la loi de S.

<u>SOLUTION</u>:

Remarquons d'abord que $Supp(S) = \mathbb{R}^+$. Soient $x \geq 0,\ f_S$ la fonction de densité de S, on a par la proposition précédente :

$$\begin{aligned} f_S(x) &= \int_{\mathbb{R}} f_X(x-y)f_Y(y)dy \\ &= \int_0^x e^{-(x-y)}e^{-y}dy \\ &= xe^{-x}. \end{aligned}$$

Ainsi

$$f_S(x) = xe^{-x}\mathbb{1}_{[0,+\infty[}(x).$$

5.5 Exercices

Exercice 5.1. Soit F la fonction définie par

$$F(x,y) = \mathbb{1}_{[1,+\infty[}(x+y).$$

Montrer que F n'est pas une fonction de répartition.

Exercice 5.2. Soient $\beta > 0, X, Y$ deux v. a. r discrètes avec

$$P_{(X,Y)}(i,j) = e^{-1}\frac{(i+j)\beta^{i+j}}{i!\ j!}$$

1. Détermier β pour que $P_{(X,Y)}$ soit une fonction de masse.
2. Calculer les lois marginales p_X et p_Y.
3. Calculer $\mathbb{E}(2^{X+Y})$.

Exercice 5.3. Soit le couple aléatoire (X,Y) de fonction de densité :

$$f_{(X,Y)}(x,y) = e^{-x-y}\mathbb{1}_{\{x>0,\ y>0\}}.$$

Trouver

1. Les densités marginales de X et de Y.
2. Les fonctions de répartitions : $F_{(X,Y)},\ F_X$ et F_Y.
3. $\mathbb{E}(X),\ \mathbb{E}(Y)$ et $Cov(X,Y)$.

Exercice 5.4. Soit (X,Y) un couple aléatoire de fonction de densité

$$f_{(X,Y)}(x,y) = e^{-y}\mathbb{1}_{\mathbb{R}_+\times\ [x,+\infty[}(x,y).$$

1. Vérifier que f est bien une densité de probabilité.
2. Donner les lois marginales de X et Y.
3. Calculer $P(X \leq 1/Y > 2)$.
4. Déterminer $f_{X/Y=y}$ et $f_{Y/X=x}$.
5. En déduire $\mathbb{E}(X/Y=y)$ et $\mathbb{E}(Y/X=x)$.
6. Calculer $\mathbb{E}(X), \mathbb{E}(Y)$ par deux méthodes différentes.

Exercice 5.5. Soient X, Y deux v. a. r avec

$$f_Y(y) = \frac{1}{y^2}\mathbb{1}_{[1,+\infty[}, \qquad f_{X/Y=y}(x) = xy^2 e^{-xy}\mathbb{1}_{\mathbb{R}_+}(x).$$

1. Donner la loi du couple (X, Y).
2. En déduire la loi marginale de X.
3. Calculer $f_{Y/X=x}$.
4. En déduire $\mathbb{E}(Y/X)$.

Exercice 5.6. Soit (X, Y) un couple aléatoire de fonction de densité

$$f_{(X,Y)}(x, y) = \frac{x}{\sqrt{2\pi}} e^{\frac{-x(x+y)}{2}} \mathbb{1}_{\mathbb{R}_+^* \times \mathbb{R}_+^*}(x, y).$$

1. Déterminer $f_{X/Y=y}$ et $f_{Y/X=x}$.
2. En déduire $\mathbb{E}(X/Y = y)$ et $\mathbb{E}(Y/X = x)$.

Exercice 5.7. Soient $n \geq 1$, $(U_k)_{1\leq k\leq n}$ une suite de variables alétoires indépendantes de même loi uniforme sur $[0, 1]$. Posons $Z_n = n \min(U_1, \ldots, U_n)$. Trouver la loi de Z_n.

Exercice 5.8. Soit (X, Y) un couple aléatoire sur
$mathbbR^2$, dont la loi admet une densité f donnée par

$$f(x, y) = n(n-1)(y-x)^{n-2}\mathbb{1}_{(0\leq x\leq y\leq 1)}.$$

où $n \geq 2$.

a. Quelle est la loi de X ?

b. Donner $f_{Y/X=x}$.

c. En déduire $\mathbb{E}(Y/X = x)$ et $\mathbb{E}(Y/X)$.

Exercice 5.9. Soient X, Y deux de v.a r, pour $n, k \geq 1$, on a :

$$P(Y = n) = \frac{n+1}{2^{n+2}}$$

et

$$P(X = k/Y = n) = \frac{1}{n+1} \text{ pour } 0 \leq k \leq n$$

1. Donner la loi du couple (X, Y) et calculer $P(X \leq Y)$.
2. Les variables aléatoires X et $X - Y$ sont-elles indépendantes ?

Exercice 5.10. Soit T une v. a. r ayant une densité de probabilité f continue sur $[0, \infty[$. Montrer que T a une loi exponentielle ssi pour tout $t > 0$, la loi conditionnelle de $T - t$ sachant $T > t$ est encore celle de T.

Exercice 5.11. Soient $\lambda,\ \mu \in \mathbb{R}^+,\ (X, Y)$ un couple aléatoire de fonction de densité

$$f(x, y) = \lambda\mu e^{-(\lambda x+\mu y)} \mathbb{1}_{[0,+\infty[}(x, y).$$

– Calculer les densités marginales.
– En déduire $\mathbb{E}\Big(X/Y = y\Big)$ et $\mathbb{E}\Big(Y/X = x\Big)$.
– Calculer $\mathbb{E}(X), \mathbb{E}(X),\ Cov(X, Y)$ et $Var(X + Y)$.

Exercice 5.12. Soient $0 < p < 1, n \geq 1, m \geq 1$ et soient X et Y deux variables aléatoires indépendantes avec $X \hookrightarrow \mathcal{B}(n, p),\ Y \hookrightarrow \mathcal{B}(m, p)$. On pose $S = X + Y$. Montrer que la variable $S \hookrightarrow \mathcal{B}(n + m, p)$.

Exercice 5.13. Soient X, Y deux v. a. r indépendantes de même loi géométrique de paramètre $p \in]0, 1[$. i. e

$$P(X = k) = p(1 - p)^k, \qquad k \in \mathbb{N}.$$

On pose $S = X + Y$

1. Donner la loi de S.
2. Calculer par deux méthodes différentes $\mathbb{E}(S)$.
3. Déterminer $\mathbb{E}(X/S)$ par deux méthodes différentes.

Exercice 5.14. Soit (X, Y) un couple aléatoire de fonction de densité

$$f(x, y) = 3x^2y(1 - y)e^{-x}\mathbb{1}_{]0,+\infty[\times]0,1[}(x, y).$$

1. Calculer les densités marginales du couple (X, Y).

2. Les variables X et Y sont-elle indépendantes ?
3. Calculer $\mathbb{E}(X), \mathbb{E}(Y)$ et $\mathbb{E}(XY)$.

Exercice 5.15. Soient (X, Y) un couple aléatoire de fonction de densité

$$f(x,y) = \frac{2}{(1+x+y)^3} \mathbb{1}\Big(x > 0, y > 0\Big).$$

1. Calculer laa fonction de répartition du couple (X, Y).
2. Déterminer les lois marginales.
3. Calculer $f_{Y/X=x}$.

Chapitre 6

Convergence de suites de variables aléatoires

Tout au long de ce chapitre, $\{X_n,\ n \geq 1\}$ désignera une suite de variables aléatoires réelles définies sur un espace de probabilité $(\Omega, \mathcal{F}, I\!P)$.

6.1 Convergence en probabilité

Définition 6.1. On dit qu'une suite de variables aléatoires $\{X_n,\ n \geq 1\}$ converge en probabilité vers une variable aléatoire X, si

$$\forall \varepsilon > 0, \qquad \lim_{n \to +\infty} I\!P(|X_n - X| \geq \varepsilon) = 0.$$

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

Exemple 6.1. Considérons la suite de variables aléatoires $\{X_n,\ n \geq 1\}$ où chaque variable X_n prend les valeurs 0 et 1 avec les probabilités respectives $1 - \frac{1}{n}$ et $\frac{1}{n}$, alors la suite $\{X_n,\ n \geq 1\}$ converge vers 0 en probabilité. En effet si $\varepsilon > 0$, alors deux cas peuvent se présenter :

Premier cas : $\varepsilon > 1$.

Dans ce cas, il est clair que pour tout $n \geq 1$,

$$I\!P(|X_n - 0| \geq \varepsilon) = I\!P(X_n \geq \varepsilon) = 0.$$

DEUXIÈME CAS : $0 \leq \varepsilon \leq 1$.

Dans ce cas, on a

$$\mathbb{P}(|X_n - 0| \geq \varepsilon) = \mathbb{P}(X_n \geq \varepsilon) = \mathbb{P}(X_n = 1) = \frac{1}{n},$$

d'où $X_n \xrightarrow{P} 0$.

Proposition 6.1. Soient $\{X_n,\ Y_n\ n \geq 1\}$ deux suites de variables aléatoires, $f:\ \mathbb{R} \to \mathbb{R}$ une fonction continue. Supposons que $\{X_n,\ n \geq 1\}$ (resp. $\{Y_n\ n \geq 1\}$) converge en probabilité vers une variable aléatoire X (resp. Y), alors

- Pour tout $\alpha,\ \beta \in \mathbb{R},\ \alpha X_n + \beta Y_n \xrightarrow{P} \alpha X + \beta Y$.
- $X_n Y_n \xrightarrow{P} XY$.
- $f(X_n) \xrightarrow{P} f(X)$.

6.2 Convergence en moyenne

Définition 6.2. Soit X une variable aléatoire réelle.

- On dit que X est intégrable si $\mathbb{E}(|X| < \infty)$.
- On dit que X est de carré intégrable si $\mathbb{E}(|X|^2 < \infty)$.
- Si $p \geq 1$, on dit que X est dans $L^p(\Omega)$ si $\mathbb{E}(|X|^p < \infty)$.

Définition 6.3. On dit qu'une suite de variables aléatoires intégrables $\{X_n,\ n \geq 1\}$ converge en moyenne vers une variable aléatoire X, si

$$\lim_{n \to +\infty} \mathbb{E}(|X_n - X|) = 0.$$

On note $X_n \to X$ en moyenne.

Définition 6.4. On dit qu'une suite de variables aléatoires de carré intégrable $\{X_n,\ n \geq 1\}$ converge en moyenne quadratique vers une variable aléatoire X, si

$$\lim_{n \to +\infty} \mathbb{E}(|X_n - X|^2) = 0.$$

On note $X_n \to X$ en moyenne quadratique.

Remarque 6.1. Il faut bien noter que

– On ne peut pas parler de la convergence en moyenne (resp. en moyenne quadratique) si les variables considérées ne sont pas intégrables (resp. de carré intégrable).

– Lorsqu'une suite de variables aléatoires intégrables $\{X_n,\ n \geq 1\}$ converge en moyenne (resp. en moyenne quadratique) vers une variable aléatoire X, alors X est aussi intégrable i. e. $\mathbb{E}(|X| < \infty)$.(resp. de carré intégrable i. e. $\mathbb{E}(|X|^2) < \infty$).

Exemple 6.2. Considérons la suite de variables aléatoires $\{X_n,\ n \geq 1\}$ où chaque variable X_n prends les valeurs 0 et 1 avec les probabilités respectives $1 - e^{-n}$ et e^{-n}, alors la suite $\{X_n,\ n \geq 1\}$ converge vers 0 en moyenne. En effet si $n \geq 1$, alors

Dans ce cas, il est clair que pour tout $n \geq 1$,

$$\mathbb{E}(|X_n - 0|) = \mathbb{E}(X_n) = e^{-n} \longrightarrow 0,$$

lorsque $n \to +\infty$.

6.3 Convergence presque sûre

Définition 6.5. On dit qu'une suite de variables aléatoires $\{X_n,\ n \geq 1\}$ converge presque sûrement vers une variable aléatoire X, si

$$I\!P\{\omega \in \Omega,\ \lim_{n \to +\infty} X_n(\omega) = X(\omega)\} = 1$$

On note $X_n \longrightarrow X$ p. s.

Un critère très utilisé pour établir la convergence presque sûre pour une suite des variables aléatoires réelles est donné par le résultat suivant :

Théorème 6.1. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires.

– Supposons que pour tout $\varepsilon > 0$,

$$\sum_{n \geq 1} I\!P\Big(|X_n - X| \geq \varepsilon\Big) < \infty, \tag{6.1}$$

alors la suite $\{X_n,\ n \geq 1\}$ converge presque surement vers X.

– Réciproquement, si $\{X_n,\ n \geq 1\}$ est une suite de v. a. r indépendantes qui converge presque sûrement vers X, alors la condition 6.1 est vérifiée.

Exemple 6.3. Considérons la suite de variables aléatoires $\{X_n,\ n \geq 1\}$ où chaque variable X_n prend les valeurs 0 et 1 avec les probabilités respectives $1 - \frac{1}{n^2}$ et $\frac{1}{n^2}$, alors la suite $\{X_n,\ n \geq 1\}$ converge vers 0 presque sûrement. En effet si $\varepsilon > 0$, alors on peut considérer les deux cas suivants :

PREMIER CAS : $\varepsilon > 1$.

Il est clair que dans ce cas, on a pour tout $n \geq 1$,

$$I\!P(|X_n - 0| \geq \varepsilon) = I\!P(X_n \geq \varepsilon) = 0.$$

DEUXIÈME CAS : $0 \leq \varepsilon \leq 1$.

Dans ce cas, on a

$$I\!P(|X_n - 0| \geq \varepsilon) = I\!P(X_n \geq \varepsilon) = I\!P(X_n = 1) = \frac{1}{n^2}.$$

Comme $\sum_{n\geq 1} \frac{1}{n^2} < \infty$, on a bien

$$\sum_{n\geq 1} I\!P(|X_n - 0| \geq \varepsilon) < \infty.$$

Ainsi la condition 6.1 est vérifiée et la suite $\{X_n,\ n \geq 1\}$ converge presque sûrement vers 0.

6.4 Convergence en loi

On rappelle que si X est une variable aléatoire, sa loi que nous avons notée $I\!P_X$ est une probabilité définie sur $(\mathbb{R},\ \mathcal{B}_{\mathbb{R}})$ par

$$\forall A \in \mathcal{B}_{\mathbb{R}}, \qquad I\!P_X(A) = I\!P(X \in A).$$

On rappelle aussi que si X est une variable aléatoire réelle, on peut lui associer une fonction réelle qu'on appelle fonction de répartition, définie par

$$\forall x \in \mathbb{R}, \qquad F_X(x) = I\!P(X \leq x)$$

et une autre fonction appelée fonction caractéristique et notée φ_X définie sur $\mathbb{R}$ par

$$\forall t \in \mathbb{R}, \qquad \varphi_X(t) = \mathbb{E}\Big(e^{itX}\Big).$$

On a vu aussi que deux variables aléatoires ont même loi (i. e. $I\!P_X = I\!P_Y$) si et seulement si elles ont la même fonction de répartition i. e. $F_X = F_Y$.

Pour une suite de variables aléatoire $\{X_n,\ n \geq 1\}$, on note $\{F_n,\ n \geq 1\}$ (resp. φ_n) la suite des fonctions de répartitions associées (resp. la suite des fonctions caractéristiques associées) i.e. pour tout $x,\ t \in \mathbb{R}$:

$$\forall n \geq 1, \qquad F_n(x) = F_{X_n}(x) \quad \text{ et } \quad \varphi_n(t) = \varphi_{X_n}(t).$$

Définition 6.6. Soient $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles, $\{F_n,\ n \geq 1\}$ la suite des fonctions de répartitions correspondantes. On dit que $\{X_n,\ n \geq 1\}$ converge en loi vers une variable aléatoire X de fonction de répartition F_X, si

$$\lim_{n \to +\infty} F_n(t) = F_X(t),$$

pour tout point t où F_X est continue. On note $X_n \xrightarrow{\mathcal{L}} X$.

Exemple 6.4. Soient $\alpha > 0$, $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires indépendantes de même fonction de densité :

$$f(x) = \alpha x^{-(\alpha+1)} \mathbb{1}_{]1,+\infty[}(x).$$

et posons pour $n \geq 1$:

$$Y_n = n^{-\frac{1}{\alpha}} \max_{1 \leq k \leq n} X_k.$$

La suite de variables aléatoires $\{Y_n,\ n \geq 1\}$ converge en loi vers la variable Y de loi

$$f_Y(y) = \alpha y^{-(\alpha+1)} e^{-y^{-\alpha}} \mathbb{1}_{]0,+\infty[}(y).$$

En effet, on a pour tout $x > 1$, la fonction de répartition pour chaque variable X_n est

$$F(x) = \int_0^x \alpha t^{-(\alpha+1)} dt = 1 - x^{-\alpha}.$$

Par suite, si $x > 0$ on peut écrire :

$$\begin{aligned}
F_{Y_n}(x) &= I\!P\Big(n^{-\frac{1}{\alpha}} \max_{1\leq k\leq n} X_k \leq x\Big) \\
&= I\!P\Big(\max_{1\leq k\leq n} X_k \leq n^{\frac{1}{\alpha}} x\Big) \\
&= I\!P\Big(\forall 1 \leq k \leq n,\ X_k \leq n^{\frac{1}{\alpha}} x\Big) \\
&= \Big(F\Big(n^{\frac{1}{\alpha}} x\Big)\Big)^n \\
&= \Big(1 - \frac{1}{n x^{\alpha}}\Big)^n.
\end{aligned}$$

Par conséquent,

$$\forall x > 0, \qquad F_{Y_n}(x) \longrightarrow e^{-x^{-\alpha}},$$

lorsque $n \to +\infty$, pour tout $x > 0$. Ceci prouve la convergence en loi de la suite $\{Y_n,\ n \geq 1\}$. En dérivant cette dernière fonction on obtient la loi de Y.

Dans bien des situations, il est difficile de montrer la convergence en loi d'une suite da variables aléatoires en utilisant la définition, c'est pourquoi le résultat suivant peut être d'une grande utilité :

Proposition 6.2. Soient $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles, $\{\varphi_n,\ n \geq 1\}$ la suite des fonctions caractéristiques correspondantes. Les trois propriétés suivantes sont équivalentes :

- $\{X_n,\ n \geq 1\}$ converge en loi vers X.
- Pour tout $t \in \mathbb{R}$, $\quad \lim_{n\to+\infty} \varphi_n(t) = \varphi_X(t)$.
- Pour toute fonction continue et bornée $f :\ \mathbb{R}\ \to\ \mathbb{R}$:

$$\lim_{n\to+\infty} \mathbb{E}\Big(f(X_n)\Big) = \mathbb{E}\Big(f(X)\Big).$$

De cette dernière proposition, on peut déduire le résultat suivant :

Proposition 6.3. Soient $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles, $h : \mathbb{R} \to \mathbb{R}$ une fonction <u>continue</u>. Si la suite $\{X_n,\ n \geq 1\}$ converge en loi vers une variable aléatoire X alors la suite $\{h(X_n),\ n \geq 1\}$ converge en loi vers la variable aléatoire $h(X)$.

Exemple 6.5. Soient $\{p_n,\ n \geq 1\}$ une suite de nombres réels, $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles. Supposons que pour tout $n \geq 1$:

$$0 < p_n < 1, \qquad X_n \hookrightarrow \mathcal{B}(n, p_n)$$

et qu'il existe $\lambda > 0$, vérifiant

$$\lim_{n \to +\infty} np_n = \lambda.$$

Alors il existe une variable aléatoire $X \hookrightarrow \mathcal{P}(\lambda)$, telle que $X_n \xrightarrow{\mathcal{L}} X$.

<u>Solution</u>:

Faisont appel au deuxième point de la Proposition 6.2 pour établir ce résultat, en effet on sait (*voir Chapitre 4*) que si $X \hookrightarrow \mathcal{P}(\lambda)$, alors pour tout $t \in \mathbb{R}$,

$$\varphi_X(t) = e^{\lambda(e^{it}-1)}.$$

De même, on a vu au chapitre 4 aussi que si φ_n est la fonction caractéristique de la variable X_n, alors on a pour $t \in \mathbb{R}$,

$$\varphi_n(t) = \Big(1 - p_n + p_n e^{it}\Big)^n.$$

Par conséquent

$$\begin{aligned} \varphi_n(t) &= \Big(1 - p_n(1 - e^{it})\Big)^n \\ &= \Big(1 - np_n \frac{(1 - e^{it})}{n}\Big)^n. \end{aligned}$$

Mais

$$\lim_{n \to +\infty} np_n = \lambda$$

Ainsi on a

$$\lim_{n\to+\infty}\varphi_n(t)=\lim_{n\to+\infty}\Big(1-np_n\frac{(1-e^{it})}{n}\Big)^n=e^{\lambda(e^{it}-1)}=\varphi_X(t).$$

D'où le résultat.

6.5 Comparaison des modes de convergence

On a défini quatre modes de convergence, se sont les plus imprtants en probabilité, à présent nous étudions les relations qui existent entre ces modes de convergence, afin de comprendre les différentes situations qui peuvent se présenter. Nous commençons par le lien entre la convergence presque sûre et la convergence en probabilité :

Proposition 6.4. Soit $\{X_n,\ n\geq 1\}$ une suite de variables aléatoires réelles, alors

– Si $\{X_n,\ n\geq 1\}$ converge presque sûrement vers X alors elle converge en probabilité vers X.
– Il existe une suite $\{X_n,\ n\geq 1\}$ qui converge en probabilité mais qui ne converge pas presque sûrement.
– Si $\{X_n,\ n\geq 1\}$ converge en probabilité vers X alors il existe une suite de nombres naturels positifs $\{m_n,\ n\geq 1\}$ strcictement croissante, telle que la sous-suite $\{X_{m_n},\ n\geq 1\}$ converge presque sûrement vers X.

Pour la convergence en moyenne et en moyenne quadratique, on a

Proposition 6.5. Soit $\{X_n,\ n\geq 1\}$ une suite de variables aléatoires réelles.

– Si $\{X_n,\ n\geq 1\}$ converge en moyenne vers X alors elle converge en probabilité vers X.
– Il existe une suite $\{X_n,\ n\geq 1\}$ qui converge en probabilité mais qui ne converge pas en moyenne.
– Si $\{X_n,\ n\geq 1\}$ converge en moyenne quadratique vers X alors elle converge en moyenne vers X et donc en probabilité vers X.

En fin, concernant la convergence en loi, nous avons

Proposition 6.6. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles, alors

- Si $\{X_n,\ n \geq 1\}$ converge en probabilité vers une v. a. r X, elle converge aussi en loi vers X.
- Il existe une suite $\{X_n,\ n \geq 1\}$ qui converge en loi mais qui ne converge pas en probabilité.
- Si $\{X_n,\ n \geq 1\}$ converge en loi vers une constante C alors elle converge aussi en probabilité vers C.

Nous n'allons pas donner la preuve des liens entre ces différents modes de convergence, car elle nécesssite des outils mathématiques qui n'entrent pas dans le cadre de ce cours (théorie de la mesure), cependant nous donnerons quelques exemples simples afin d'illustrer le fait qu'on peut avoir la convergence dans un mode précis et pas dans un autre.

La convegence en probabilité n'implique pas la convergence presque sûre :

Soit $\{X_n,\ n \geq 1\}$ la suite de variables aléatoires indépendantes où chaque variable X_n prend les valeurs 0 et 1 avec les probabilité respectives $1-\frac{1}{n}$ et $\frac{1}{n}$. On a vu dans l'exemple 6.1 que cette suite converge en probabilité vers 0, cependant si $0 < \varepsilon < 1$, alors

$$\sum_{n\geq 1} I\!P(|X_n - 0| > \varepsilon) = \sum_{n\geq 1} I\!P(X_n = 1) = \sum_{n\geq 1} \frac{1}{n} = +\infty.$$

D'après le deuxiemme point du Théorème 6.1, la suite $\{X_n,\ n \geq 1\}$ ne peut pas converger vers 0 presque sûrement.

La convegence en probabilité n'implique pas la convergence en moyenne :

Soit $\{X_n,\ n \geq 1\}$ la suite de variables aléatoires où chaque variable X_n prend les valeurs 0 et n avec les probabilité respectives $1-\frac{1}{n}$ et $\frac{1}{n}$. Comme pour l'exemple 6.1, il n'est pas difficile de vérifier que cette suite converge en probabilité vers 0. D'autre part

$$\mathbb{E}(|X_n - 0|) = \mathbb{E}(X_n) = 1.$$

Ainsi cette suite ne converge pas en moyenne.

La convegence en loi n'implique pas la convergence en probabilité :

Soit $X \hookrightarrow \mathcal{N}(0,1)$, $Y = -X$ et posons pour $n \geq 1$, $X_n = X$. Une variable de loi $\mathcal{N}(0,1)$ est symétrique, ainsi X et Y ont la même loi, d'autre part il est évident que X_n converge en loi vers X donc vers Y aussi. Mais comme $Z := X - Y \hookrightarrow \mathcal{N}(0,2)$, on peut voir que

$$\mathbb{P}(|X_n - Y| \geq 2) = \mathbb{P}(|X - Y| \geq 2) = \mathbb{P}(|Z| \geq 2) = \frac{1}{\sqrt{2\pi}} \int_1^{+\infty} e^{-\frac{x^2}{8}} dx > 0.$$

Ainsi, la suite $\{X_n,\ n \geq 1\}$ ne converge pas en probabilité.

La convegence presque sûre n'implique pas la convergence en moyenne :

Soit $\{X_n,\ n \geq 1\}$ la suite de variables aléatoires où chaque variable X_n prend les valeurs 0 et n^2 avec les probabilités respectives $1 - \frac{1}{n^2}$ et $\frac{1}{n^2}$. Remarquons tout d'abord que pour tout $\varepsilon > 0$,

$$\sum_{n \geq 1} \mathbb{P}(|X_n - 0| > \varepsilon) = \sum_{n \geq 1} \mathbb{P}(X_n = n^2) = \sum_{n \geq 1} \frac{1}{n^2} < \infty.$$

A présent, le Théorème 6.1 permet de conclure que la suite $\{X_n,\ n \geq 1\}$ converge presque sûrement vers 0. D'un autre coté

$$\mathbb{E}(|X_n - 0|) = \mathbb{E}(X_n) = 1.$$

Cette suite ne peut pas converger vers 0 en moyenne.

La convegence en moyenne n'implique pas la convergence presque sûre :

Soit $\{X_n,\ n \geq 1\}$ la suite de variables aléatoires indépendantes où chaque variable X_n prend les valeurs 0 et $\sqrt{n}$ avec les probabilités respectives $1 - \frac{1}{n}$ et $\frac{1}{n}$. En faisant appel encore une fois au Théorème 6.1, on montre que cette suite ne converge pas presque sûrement. D'autre part

$$\mathbb{E}(|X_n - 0|) = \mathbb{E}(X_n) = \frac{1}{\sqrt{n}} \longrightarrow 0.$$

Par suite, cette suite converge vers 0 en moyenne. On arrive donc au résultat suivant :

Proposition 6.7. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires réelles, alors

- La convergence presque sûre implique la convergence en probabilité.
- La convergence en probabilité implique la convergence en loi.
- La convergence en moyenne implique la convergence en probabilité.
- Les implications réciproques sont en général fausses.

6.6 Lois des grands nombres

Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoires indépendantes, Dans ce paragraphe on s'intéresse au comportement des moyennes arithmétiques

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k,$$

lorsque n devient de plus en plus grands. Les résultats concernant ce problème sont appelés : Lois des Grands Nombres : (LGN). Ces résultats ce décomposent en deux parties :

- Lois fortes des grands nombres.
- Lois faibles des grands nombres.

La différence entre ces deux familles de résultats réside dans le mode de comportement qu'on étudie, dans la première on s'intéresse au comportement ponctuel(convergence presque sûre), alors que dans la seconde on regarde le comportement en probabilité (convergence en probabilité) qui est plus faible.

Nous commençons par établir deux résultats de la seconde famille de lois :

Proposition 6.8. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoire réelles indépendantes, centrées et de carré intégrable. Supposons que

$$\frac{1}{n^2}\sum_{k=1}^{n} Var(X_k) \longrightarrow 0,$$

alors les moyennes $\{\overline{X}_n,\ n \geq 1\}$ convergent vers 0 en probabilité.

Comme corollaire direct de ce résultat, on obtient

Théorème 6.1. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoire réelles indépendantes, de même loi et de carré intégrable. Posons $\mu = \mathbb{E}(X_1)$, alors

$$\overline{X}_n \longrightarrow \mu,$$

en probabilité lorsque $n \to \infty$.

PREUVE

Pour $k \geq 1$, posons $Y_k = X_k - \mu$, $\sigma^2 = Var(X_k) = Var(Y_k)$. La suite $\{Y_n,\ n \geq 1\}$ est par construction une suite de v. a. r indépendantes et centrées, de plus

$$\frac{1}{n^2}\sum_{k=1}^{n} Var(Y_k) = \frac{\sigma^2}{n} \longrightarrow 0.$$

Par la proposition précédente,

$$\frac{1}{n}\sum_{k=1}^{n} (X_k - \mu) \longrightarrow 0,$$

en probabilité, ainsi

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow \mu,$$

en probabilité.

Pour la seconde famille des lois des grands nombres, on a la loi forte des grands nombres de Kolmogorov :

Théorème 6.2. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoire réelles, indépendantes, intégrables et de même loi. Posons $\mu = \mathbb{E}(X_1)$, alors

$$\overline{X}_n \longrightarrow \mu,$$

presque surement lorsque $n \to \infty$.

6.7 Le théorème de la limite centrale

Le théorème de la limite centrale (TLC) est un théorème fondamental, non seulement en calcul des probabilités mais également en statistique mathématique, il affirme grosso modo qu'une somme finie de variables aléatoires indépendantes, centrées, de même loi et de variance finie, se comporte en loi (lorsqu'elle est bien normalisée et lorsque n devient très grand) comme une variable aléatoire de loi normale centrée réduite. Ainsi on a pas besoin de connaitre la loi des variables aléatoire considérées. Ce résultat permet (entre autre) de construire un intervalle de confiance pour un paramètre inconnu, ou alors de faire un test d'hypothèse dans la statistique mathématique.

Théorème 6.3. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoire réelles, indépendantes, de même loi et de carré intégrable. Posons $\mu = \mathbb{E}(X_1)$ et
$\sigma^2 = Var(X_1)$, alors

$$\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1),$$

lorsque $n \to +\infty$. Par conséquent

$$\forall x \in \mathbb{R}, \qquad \lim_{n\to+\infty} I\!P\Big(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu) \leq x\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Remarque 6.2. Il faut bien noter qu'on n'exige aucune information sur la loi des variables aléatoires conidérées.

<u>PREUVE</u>

Pour $k \geq 1$, posons

$$Y_k = \frac{X_k - \mu}{\sigma}.$$

Les variables Y_k ainsi définies sont indépendantes, de même loi, centrées et réduites i.e. $\mathbb{E}(Y_k) = 0$ et $Var(Y_k) = 1$. Posons pour tout $k \geq 1,\ t \in \mathbb{R}$:

$$Z_k = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} Y_k,$$

$$\varphi_k(t) = \mathbb{E}\Big(e^{itZ_k}\Big) \qquad \text{et} \qquad \varphi(t) = \mathbb{E}\Big(e^{itY_1}\Big).$$

Soit à présent $n \geq 1$. En faisant appel à l'indépendance des variables Y_k, on peut écrire

$$\begin{aligned}
\varphi_n(t) &= \mathbb{E}\Big(e^{\frac{it}{\sqrt{n}}\sum_{k=1}^n Y_k}\Big) \\
&= \prod_{k=1}^n \mathbb{E}\Big(e^{\frac{it}{\sqrt{n}}Y_k}\Big) \\
&= \prod_{k=1}^n \varphi(\frac{tY_1}{\sqrt{n}}) \\
&= \varphi^n(\frac{tY_1}{\sqrt{n}}).
\end{aligned}$$

D'autre part, les Y_k ont une moyenne nulle et une variance qui vaut 1, par la proposition 3.16 , la fonction φ admet au voisinage de 0, le dévelopement limité suivant :

$$\varphi(u) = 1 - \frac{u^2}{2} + o(u^2).$$

Donc pour n assez grand on a

$$\varphi(\frac{tY_1}{\sqrt{n}}) = 1 - \frac{t^2}{2n} + o(\frac{t^2}{2n}).$$

Ainsi

$$\lim_{n\to+\infty} \varphi_n(t) = \lim_{n\to+\infty} \varphi^n(t) = \lim_{n\to+\infty} \Big(1 - \frac{t^2}{2n} + o(\frac{t^2}{2n})\Big)^n = e^{-\frac{t^2}{2}}.$$

La proposition 6.2 permet de conclure que la suite $\{Z_n,\ n \geq 1\}$ converge en loi vers une variable aléatoire de loi $\mathcal{N}(0,1)$, ceci prouve notre théorème.

En combinant ce théorème avec la proposition 6.3 on obtient

Proposition 6.9. Soit $\{X_n,\ n \geq 1\}$ une suite de variables aléatoire réelles, indépendantes, de même loi et de carré intégrable. Posons $\mu = \mathbb{E}(X_1)$ et $\sigma^2 = Var(X_1)$, alors

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n X_k \xrightarrow{\mathcal{L}} \mathcal{N}(\mu, \sigma^2),$$

lorsque $n \to +\infty$. Par conséquent

$$\forall x \in \mathbb{R}, \qquad \lim_{n \to +\infty} \mathbb{P}\Big(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k \leq x\Big) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt.$$

6.8 Exercices

Exercice 6.1. Soient $0 < a < 1$, $X \hookrightarrow \mathcal{G}(a)$.

1. Calculer F_X et φ_X.

2. Calculer $\mathbb{E}(X)$ et $Var(X)$ par deux méthodes différentes.

A présent, Soit $(X_n)_{n\geq 1}$ une suite de v. a. r avec $X_n \hookrightarrow \mathcal{G}(\frac{a}{n})$.

Montrer par deux méthodes différentes que $(\frac{X_n}{n})_{n\geq 1}$ converge en loi vers une v. a. X à déterminer.

Exercice 6.2. Soient $0 < a < 1$, $(X_n)_{n\geq 1}$ une suite de v. a. r.

1. On suppose que

$$\forall n \geq 1, \qquad X_n \hookrightarrow \mathcal{U}\{1, 2, \ldots, n\}.$$

Montrer que $(\frac{X_n}{n})_{n\geq 1}$ converge en loi vers une v. a. U à déterminer.

2. On suppose que

$$\forall n \geq 1 \qquad X_n \hookrightarrow \mathcal{B}(n, \frac{a}{n}).$$

Montrer que $(X_n)_{n\geq 1}$ converge en loi vers une v. a. X à déterminer.

Exercice 6.3. La probabilité qu'une pièce de monnaie montre pile lors d'un lancers est $p \in]0, 1[$. On lance cette pièce n fois et on note respectivement P_n et F_n le nombre de piles et de faces obtenus lors de ces lancers qu'on suppose indépendants.

1. Quelle est la loi de la v. a. r $P_n + F_n$?

2. Expliquer pourquoi les suites $(\frac{P_n}{n})_{n\geq 1}$ et $(\frac{F_n}{n})_{n\geq 1}$ converge presque sûrement et donner leurs limites.

3. Montrer que

$$P\Big(2p - 1 - \varepsilon \leq \frac{1}{n}(P_n - F_n) \leq 2p - 1 + \varepsilon\Big) \longrightarrow 1,$$

lorsque $n \longrightarrow \infty$.

4. Montrer que pour tout $n \in \mathbb{N}^*$,

$$\mathbb{E}(\frac{1}{1 + P_n}) = \frac{1 - (1-p)^{n+1}}{(n+1)p}.$$

Exercice 6.4. Soit $(X_n)_{n\geq 1}$ une suite de v. a. indépendantes de même loi $\mathcal{P}(1)$. On pose $S_n = \sum_1^n X_k$.

1. Montrer par récurrence que $S_n \hookrightarrow \mathcal{P}(n)$, pour tout $n \geq 1$.
2. Appliquer le TLC à la suite $(X_n)_{n\geq 1}$.
3. En déduire que
$$\lim_{n\to\infty} e^{-n} \sum_{k=1}^{n} \frac{n^k}{k!} = \frac{1}{2}.$$

Exercice 6.5. Soient $(U_n)_{n\geq 1}$ une suite de v. a. r indépendantes de loi uniforme sur $[0,1]$. On pose pour tout $n \geq 1$:

$$Y_n = e^{\alpha\sqrt{n}} \Big(\prod_{k=1}^{n} U_k \Big)^{\frac{a}{\sqrt{n}}}$$

1. Etudier la convergence en loi de $Z_n = \ln(Y_n)$.
2. Etudier la convergence en loi de Y_n.

Exercice 6.6. Soient $n \geq 1, \lambda > 0$ et $(X_k)_{k\geq 1}$ une suite de v. a. r indépendantes de même loi $\mathcal{P}(\lambda)$, posons

$$Y_n = \prod_{k=1}^{n} (1 + X_k), \qquad Z_n = \prod_{k=1}^{n} X_k.$$

1. Etudier la convergence presque sûre de $\frac{1}{n}\log(Y_n)$.
2. Calculer $P(Z_n \neq 0)$.
3. Etudier la convergence en moyenne de Z_n.

Bibliographie

[1] Allab. K., *Eléments d'analyse : fonction d'une variable réelle.* OPU. 1991.

[2] Ash. R, B., *Probability and measure theory.* Harcourt Academic Press. 2000.

[3] Barbe. P, Ledoux. M., *Probabilité.* EDP Sciences., 2007.

[4] Bertsekas. P, Tsitsiklis. J. N., *Introduction to Probability.* Course 6.041-6.431 M.I.T. FALL. 2000.

[5] Cacoullos. T., *Exercises in Probability.* Springer-Verlag. 1989.

[6] Calot. G., *Cours de calcul des Probabilités.* Dunod. 1982.

[7] Gordon. H., *Discrete Probability.* Springer-Verlag New York, Inc. 1997.

[8] Gut. A., *Probability : a graduate course.* Springer texts in statistics. 2005.

[9] Jacod. J, Protter. P., *Probability-Essentials.* Springer-Verlag Berlin and Heidelberg Second edition. 2003

[10] Lefebvre. M., *Basic Probability theory with applications.* Springer. 2000.

[11] Miri. S. E., *Algèbe et Analyse.* Polycopié, Université Abou Bekr Belkaid. 2013.

[12] Redjdal, K. *Cours de Probabilités.* O.P.U. 1995.