2024

Probabilités 1

I. Laroussi

Université Frères Mentouri
Constantine 1
20/04/2024



Table des matières

Table des matières **i**

Introduction **iii**

1 Les Probabilités **1**

1.1 L'analyse combinatoire 1

1.2 Expériences et événements aléatoires 4

1.3 Les Probabilités 6

1.4 Exercices 17

2 Les variables aléatoires discrètes et leur lois usuelles **21**

2.1 Définition générale d'une variable aléatoire 21

2.2 Variable aléatoire discrète 22

2.3 Fonction de répartition d'une v.a. discrète 24

2.4 Les moments 25

2.5 Fonctions génératrices des moments 27

2.6 Quelques lois discrètes usuelles 28

2.7 Approximation entre variables aléatoires discrètes 40

2.8 Exercices 42

3 Variables aléatoires absolument continues **47**

3.1 Les v. a. r. absolument continus 47

3.2 Fonction de répartition 49

3.3 Les moments 51

3.4 Fonction caractéristique 52

3.5 Lois de probabilités absolument continues usuelles 55

3.6 Approximation par une loi normale 69
3.7 Exercices 72

4 Vecteurs aléatoires **73**
4.1 Définitions et propriétés 73
4.2 La fonction de répartition 77
4.3 Les moments 79
4.4 Loi conditionnelle et indépendance 81
4.5 Calcul de loi 84
4.6 Vecteur Gaussien 87
4.7 Fonction caractéristique d'un couple 90
4.8 Exercices 93

5 Convergence en loi **95**
5.1 Fonction caractéristique et somme de variables 95
5.2 Convergence en loi 97
5.3 Exercices 100

Bibliographie **103**

Introduction

La théorie des probabilités n'est pas figée et se développe pour répondre à des besoins réels. La propagation d'une épidémie ou d'une pandémie, les jeux de hasard, les files d'attentes, la fiabilité des systèmes, les télécommunications, les finances ... ont été à l'origine de certains problèmes mathématiques difficiles dont la théorie des probabilités procure des solutions intégrales ou partielles. Le résultat d'un jet de dé ou d'un scrutin est un exemple simple d'événements issus d'une expérience dont le résultat ne peut être prédit. De tels événements, dépendant du hasard, sont dits aléatoires et constituent une idée importante en théorie des probabilités. Cette dernière se détermine par le nombre de cas favorables sur le nombre de cas possibles et la solution fait souvent appel au dénombrement. Ce cours a pour but de familiariser l'étudiant avec le raisonnement probabiliste. Par rapport à un autre cours de mathématiques, il se différencie par l'ambition de modéliser certaines données réelles. Un modèle mathématiquement correct ne suffirait donc pas toujours, il faudrait encore que celui-ci coïncide aux observations. Dans certaines situations, le cadre théorique précédent est insuffisant, c'est le cas en particulier quand on s'intéresse à une mesure physique (poids, tension électrique, ...) qui prend ses valeurs dans $\mathbb{R}$, lequel n'est pas un ensemble dénombrable. Ce sont alors d'autres techniques qui sont employées et qui sont consacrées à la notion de densité de probabilité et à l'approximation par la loi normale. C'est la partie "probabilités continues".

Ce polycopié vise les étudiants de $2^{\text{ième}}$ année mathématiques. Il se divise en deux parties principales : la première concerne un rappel assez conséquent sur les probabilités avec un ensemble fondamental défini, dans sa plus grande généralité, fini ou dénombrable. C'est la partie "probabi-

lités discrètes"(chapitres 1). Le chapitre 2 rappelle une introduction aux variables aléatoires discrètes. La seconde partie, englobe les probabilités continues, les variables aléatoires continues, les vecteurs aléatoires et enfin la convergence en loi avec quelques exercices. Afin de mettre l'accent sur les problèmes fondamentaux de conditionnement et de prédiction, nous l'avons réduit à sa plus simple expression l'information théorique de ce cours. Nous avons en particulier toujours privilégié la démarche constructive à la démarche axiomatique.

1 Les Probabilités

Dans la première partie, nous allons voir l'analyse combinatoire base incontournable, dans la formation, de tous probabilistes statisticiens. Son objectif est d'étudier comment dénombrer (compter) des objets (dispositions) d'un ensemble fini. En deuxième partie on définit les fondements principaux ainsi que, le langage utilisé pour parler des expériences aléatoires. Après avoir présenté les événements résultats d'expérience aléatoire, nous allons, dans ce chapitre, essayer de quantifier ou **mesurer** la réalisation de ses événements de sorte à ne plus parler d'eux comme étant des ensembles qui peuvent ou pas se réaliser mais plutôt, comme étant une quantité ou chiffre dans $[0; 1]$ qui peut correspondre à un pourcentage en multipliant par 100%.

1.1 L'analyse combinatoire

L'idée qui construit cette discipline est de trouver un nombre entier ($n \in \mathbb{N}$) qui représente le cardinal d'une partition d'un ensemble fini. Cette partition est choisie suivant des conditions ou critères spécifiques et la finalité est de savoir combien, à partir d'un ensemble Ω de n éléments, peut-on construire de groupes composés de k éléments ($k \leq n$). Ce dernier sous-groupe occupe une position dite **disposition**. Pour faciliter la compréhension, nous allons prendre l'exemple de quatre étudiants et nous voulons savoir combien de dispositions peut-on obtenir pour choisir deux étudiants parmi ces quatre. Supposons que nos étudiants sont nommés

$A,\ B,\ C$ et D, on obtient si l'ordre de notre couple est important

$$\Omega_1 = \left\{ \begin{array}{c} (A,B);(A,C);(A,D);(B,A);(B,C);(B,D); \\ (C,A);(C,B);(C,D);(D,A);(D,B);(D,C) \end{array} \right\}.$$

Si l'ordre n'est pas important, on obtient

$$\Omega_2 = \{(A,B);(A,C);(A,D);(B,C);(B,D);(C,D)\}.$$

Nous allons mieux expliquer l'idée d'ordre. Ici on dit que le couple (A,B) est une disposition et dans le premier cas, on obtient 12 dispositions alors que dans le deuxième, on obtient 6 dispositions. Dans les parties qui vont suivre, nous allons voir est nommer cette différence ainsi qu'apprendre à calculer $card(\Omega)$ sans avoir à la définir ou à définir ses éléments.

1.1.1 Type de dispositions

Maintenant et suivant le raisonnement précédent, nous allons voir qu'il existe deux type de dispositions. La première est dite **avec répétition** et c'est le cas où dans l'exemple précédent on peut obtenir des couples du genre (A,A) ou (B,B) .... Ce cas de figure ne sera pas pris en compte dans la suite de notre cour et nous allons plus travailler dans le deuxième cas qui est **sans répétition**. Nous allons voir qu'il existe deux types de dispositions et cela même sans avoir de répétition qui sont avec ordre ou sans ordre.

Arrangement

On appelle arrangement de k éléments toute disposition **ordonnée** de k éléments pris parmi n éléments de l'ensemble fini Ω. Le nombre des arrangements de k parmi n est noté A_n^k tel que

$$A_n^k = \frac{n!}{(n-k)!}, \tag{1.1}$$

avec $1 \leq k \leq n$. L'idée, pour obtenir cette formule, est que si on suppose que nous avons k emplacements et nous voulons les occuper par k éléments d'un ensemble de n éléments, alors on obtient que le 1^{er} emplacement est

convoité par n éléments, le deuxième est convoité, obligatoirement, par $n-1$ éléments et cela du fait que nous devons respecter l'ordre et ainsi de suite. A la position k il nous reste $(n-k+1)$ élément qui convoite cet emplacement. Donc, on peut l'écrire comme

$$A_n^k = n(n-1)(n-2)\dots(n-k+2)(n-k+1).$$

qui peut s'écrire sous la forme

$$A_n^k = \frac{n(n-1)(n-2)\dots(n-k+2)(n-k+1)(n-k)(n-k-1)\dots 4.3.2.1}{(n-k)(n-k-1)\dots 4.3.2.1}$$

ce qui implique l'écriture (1.1).

Dans le cas où $n = k$, nous parlons de **permutation**, puisque nous aurons dénombré toutes les possibilités de faire asseoir n individus dans n emplacements. Ce nombre est noté P_n et est égal à $P_n = n!$. On peut vérifier cette écriture en remplaçant k par n dans la relation (1.1) et en utilisant que $0! = 1$.

Combinaison

Si dans un ensemble Ω, nous voulons tirer k éléments de ces n éléments sans respecter un ordre quelconque, on obtient une **combinaison**. Le nombre totale de toutes les combinaisons possibles de k parmi n est noté C_n^k. Pour calculer ce nombre, nous allons procéder de la même manière que pour les arrangements mais en enlevant les cas de disposition semblable. C'est-à-dire que nous allons passer par deux étapes. La première est d'arranger k éléments parmi n, ce qui nous donne A_n^k possibilités. Une fois tiré, il y a $k!$ manière de les ordonner. D'où que

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

La notation C_n^k est parfois remplacer par $\begin{pmatrix} n \\ k \end{pmatrix}$. Elles sont aussi appelées les coefficients binomiaux.

Propriétés des combinaisons

Quelques propriétés seront identifiées dans cette partie, mais d'autres existent et serons rencontrées dans la suite du chapitre.

— **La symétrie**

1. $C_n^0 = C_n^n = 1$.

2. Si $n \geq 1$ alors, $C_n^1 = C_n^{n-1} = n$.
3. Si $n \geq 2$ alors, $C_n^2 = C_n^{n-2}$.
   Par conséquent, on déduit

$$C_n^k = C_n^{n-k}.$$

— **La formule de Pascal**
Si $0 \leq k \leq n-1$ alors,

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

— **La formule du binôme de Newton**
Elle correspond à la décomposition des différents termes de la puissance $n^{\text{ième}}$ du binôme $(a+b)$.

$$\forall a, b \in \mathbb{R}, n \in \mathbb{N}, (a+b)^n = \sum_{k=0}^{n} C_n^k a^{n-k} b^k.$$

1.2 Expériences et événements aléatoires

En premier lieu, nous avons une expérience, qui reproduite dans des conditions identiques conduit à plusieurs résultats que nous ne pouvons prévoir à l'avance, est dite expérience aléatoire. L'ensemble qui contient tous les résultats possibles d'une expérience aléatoire est dit **espace d'états** (ensemble fondamental associé à l'expérience). Il est noté Ω. Un résultat possible de l'expérience est noté ω tel que $\omega \in \Omega$. Un exemple simple de ce phénomène est le jet de deux pièces de monnaie, donc $\Omega = \{PP, PF, FP, FF\}$, ici $\omega_1 = PP$. Les ω_i sont appelés des **événements élémentaires** et on

appel tout sous-ensemble $A \subset \Omega$ tel que $A \neq \omega_i$ un **événement composé** associé à l'expérience aléatoire. Par exemple, A est l'événement $\ll$ le jet des pièces donne au moins un pile $\gg$, on a

$$A = \{PP, PF, FP\} = \{PP\} \cup \{PF\} \cup \{FP\}.$$

D'où l'appellation composé. D'autres cas de figure d'expériences aléatoires donnent des ensembles fondamentaux de cardinal non fini (card $\Omega = \infty$).

Donc les événements aléatoires sont des sous-ensembles de Ω et toutes les opérations élémentaires sur les ensembles seront renommées suivants le langage de l'aléatoire et cela pour décrire diverses possibilités de réalisations d'événements.

1.2.1 Relation sur les événements aléatoires

Soient deux événements aléatoires A et B.

- La réalisation de l'événement contraire "inverse" à A noté $\bar{A}$ est le résultat de l'expérience aléatoire qui n'appartient pas à A, on dit non A.
- L'événement " A **et** B se réalise en même temps", noté $A \cap B$ représente les résultats de l'expérience aléatoire appartenant à A et B en même temps.
- L'événement "A **ou** B se réalise" veut dire que l'un des deux événements est réalisé noté $A \cup B$. On dit aussi que le résultat de l'expérience se trouve dans A ou B.
- La réalisation de l'événement A **implique** la réalisation de B s'écrit $A \subset B$.
- Si l'événement $A \cap B$ donne l'ensemble vide, on dit que A et B sont deux événements **incompatibles**. On dit aussi que le résultat de l'expérience ne peut être à la fois dans A et B.
- On appel Ω l'événement **certain.** Puisque tous les résultats de l'expérience ω se trouvent dans Ω.
- Le contraire de l'événement certain ce dit l'événement **impossible** et est noté $\emptyset$.

Nous allons maintenant construire un espace $\mathcal{A}$ qui est l'ensemble de tous les événements de Ω qui vérifie des opérations ensemblistes.

1.2.2 L'algèbre et la σ-algèbre

L'algèbre est un ensemble d'ensembles qui vérifie pour $A, B \in \Omega$ alors

1. $\Omega \in \mathcal{A}$;
2. $\bar{A} \in \mathcal{A}$;
3. $A \cup B \in \mathcal{A}$.

Ce qui implique que $A \cap B \in \mathcal{A}; \emptyset \in \mathcal{A}$.
$\mathcal{A}$ une σ-algèbre est plus générale pour l'union, c-à-d,

a) $\mathcal{A}$ est une algèbre.

b) Pour une suite d'événements $(A_n)_{n\in\mathbb{N}^*}$ on ait $\cup_{n\in\mathbb{N}^*} A_n \in \mathcal{A}$.

Ce qui implique que $\mathcal{A}$ est stable par intersection dénombrable $\cap_{n\in\mathbb{N}^*} A_n \in \mathcal{A}$.

On dit que l'espace $(\Omega, \mathcal{A})$ est un espace **probabilisable.**

1.3 Les Probabilités

1.3.1 Définition et propriétés d'une probabilité

Soit $\mathbf{P}$ une application définie de $(\Omega, \mathcal{A})$ dans $[0; 1]$ tel que, pour tout événement $A \in \mathcal{A}$ on lui fait correspondre la quantité $\mathbf{P}(A)$. On écrit

$$\begin{aligned} \mathbf{P} : (\Omega, \mathcal{A}) &\longrightarrow [0; 1] \\ A &\longmapsto \mathbf{P}(A) \end{aligned}$$

Cette application vérifie les axiomes suivants

A1) $\mathbf{P}(\Omega) = 1$. L'événement certain admet une probabilité égale à un.

A2) $\mathbf{P}(A) \geq 0$ pour tout $A \in \mathcal{A}$. La probabilité de A est toujours supérieure à zéro.

A3) Si $A, B \in \mathcal{A} : A \cap B = \emptyset \Longrightarrow \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$. Si deux événements sont incompatibles alors la probabilité de leur union est égale à la somme de leurs probabilités.

A4) Si $(A_n)_{n\in\mathbb{N}^*} : A_i \cap A_j = \emptyset \ \forall i \neq j \Longrightarrow \mathbf{P}(\cup_{n\in\mathbb{N}^*} A_n) = \sum_{n\in\mathbb{N}^*} \mathbf{P}(A_n)$.

Remarque 1 *1. Si Ω est finie on a $\mathcal{A}$ est une algèbre et la probabilité est définie à l'aide des axiomes A1), A2) et A3).*

2. Si Ω est infinie dénombrable on a $\mathcal{A}$ est une σ-algèbre $\mathcal{P}(\mathbb{N})$ et la probabilité est définie à l'aide des axiomes A1), A2) et A4).

3. Si Ω est infinie (dans $\mathbb{R}$ ou bien $\Omega \subset \tau_{\mathbb{R}} = \{]-\infty,\ b[,\ b \in \mathbb{R}\}$) on a $\mathcal{A}$ est une σ-algèbre $\mathcal{B}(\mathbb{R})$ dite borélienne de $\mathbb{R}$ et la probabilité est définie à l'aide des axiomes A1), A2) et A4).

De ces trois axiomes on aboutie à quelques propriétés qui sont données par

1. $\mathbf{P}(\emptyset) = 0$, on a $\emptyset = \Omega \cap \emptyset \Longrightarrow \mathbf{P}(\Omega) = \mathbf{P}(\Omega \cup \emptyset) \overset{(A3)}{=} \mathbf{P}(\Omega) + \mathbf{P}(\emptyset)$, d'après A1) on a $\mathbf{P}(\Omega) = 1$ et $\mathbf{P}(\Omega) + \mathbf{P}(\emptyset) = 1 + \mathbf{P}(\emptyset) \Longrightarrow 1 = 1 + \mathbf{P}(\emptyset) \Longrightarrow \mathbf{P}(\emptyset) = 0$.
2. $\forall A \in \mathcal{A} : \mathbf{P}(A) = 1 - \mathbf{P}(\bar{A})$. on a

$$A \cup \bar{A} = \Omega \text{ et } A \cap \bar{A} = \emptyset, \text{ donc } \mathbf{P}(A \cup \bar{A}) \overset{(A3)}{=} \mathbf{P}(A) + \mathbf{P}(\bar{A})$$

$$\Longrightarrow \mathbf{P}(\Omega) = \mathbf{P}(A) + \mathbf{P}(\bar{A}) \overset{(A3)}{\Longrightarrow} 1 = \mathbf{P}(A) + \mathbf{P}(\bar{A}) \Longrightarrow \mathbf{P}(A) = 1 - \mathbf{P}(\bar{A}).$$

   Le contraire aussi est juste $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$.
3. $\forall A \in \mathcal{A} : 0 \leq \mathbf{P}(A) \leq 1$. D'après A2) on a $\mathbf{P}(A) \geq 0$, reste à démontrer que $\mathbf{P}(A) \leq 1$. Nous avons

$$\mathbf{P}(\Omega) = \mathbf{P}(A) + \mathbf{P}(\bar{A}) \Longrightarrow 1 = \mathbf{P}(A) + \mathbf{P}(\bar{A}) \Longrightarrow \mathbf{P}(A) \leq 1 \text{ et } \mathbf{P}(\bar{A}) \leq 1.$$

   Il existe d'autres propriétés que nous verrons à fur et à mesure.

On dit que l'espace $(\Omega, \mathcal{A}, \mathbf{P})$ est un espace **probabilisé.**

Dans la suite nous allons définir comment calculer une probabilité, à partir des axiomes et propriétés précédentes, dans le cas d'espace d'états de cardinal fini. Ce calcul va aboutir vers une propriété des plus importantes concernant l'espace des probabilités.

1.3.2 Calcul de probabilité uniforme

Soit $\Omega = \{\omega_i : i = 1, \ldots, n\}$. On parle de probabilité uniforme lorsque on est en présence d'équiprobabilité. Ce qui veut dire que les événements

élémentaires ω_i, de l'ensemble fondamental Ω (de cardinal fini), on la même probabilité de ce réaliser noté p, autrement dit

$$[\mathbf{P}(\omega_i) = p, \forall \omega_i \in \Omega].$$

Nous avons $\mathbf{P}(\Omega) = 1$ et nous avons l'écriture suivante $\Omega = \bigcup_{i=1}^{n} \omega_i$ avec ω_i incompatibles deux à deux, c'est-à-dire $\omega_i \cap \omega_j = \emptyset, \forall i \neq j$. Donc nous avons

$$\left[\mathbf{P}(\Omega) = \mathbf{P}\left(\bigcup_{i=1}^{n} \omega_i\right) = \sum_{i=1}^{n} \mathbf{P}(\omega_i) = 1\right]. \tag{1.2}$$

Et

$$\sum_{i=1}^{n} \mathbf{P}(\omega_i) = \sum_{i=1}^{n} p = np \Longrightarrow np = 1 \Longrightarrow p = \frac{1}{n}.$$

Remarque 2 *1. Dans le cas équiprobable la probabilité de l'événement élémentaire ω_i est toujours égale à $\frac{1}{n}$.*

2. D'après la relation (1.2), on obtient que dans une expérience aléatoire, quelque soit ça nature où le nombre de résultats possible est fini, la somme des probabilités des événements élémentaires est toujours égale à un.

D'après ces remarques on voit bien que la probabilité d'un événement dépend du nombre d'éléments qui le compose, donc on aboutie à la relation suivante.
Soit $A \in \mathcal{A}$ tel que $A = \bigcup_{i=1}^{k} \omega_{t_i}$ et $\{t_1, \ldots, t_k\} \subset \{i = 1, \ldots, n\}$. Pour calculer la probabilité de l'événement A nous allons utiliser l'écriture suivante

$$\mathbf{P}(A) = \mathbf{P}\left(\bigcup_{i=1}^{k} \omega_{t_i}\right) = \sum_{i=1}^{k} \mathbf{P}(\omega_{t_i}) = \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}.$$

Ainsi

$$\mathbf{P}(A) = \frac{card(A)}{card(\Omega)}.$$

Suivant ce raisonnement on obtient d'autres propriétés des probabilités.

1.3.3 Autres propriétés

Dans cette partie nous allons traiter d'autres propriétés des probabilités et la première concerne la probabilité de l'union de deux événements non incompatibles, c'est-à-dire le cas général où $A \cap B \neq \emptyset$. Nous allons avoir

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) \tag{1.3}$$

Nous avons aussi les propriétés suivantes

$$\mathbf{P}(\bar{A} \cap \bar{B}) = \mathbf{P}(\overline{A \cup B}) = 1 - \mathbf{P}(A \cup B).$$

$$\mathbf{P}(\bar{A} \cup \bar{B}) = \mathbf{P}(\overline{A \cap B}) = 1 - \mathbf{P}(A \cap B).$$

Exemple 1 *Dans une population,* 45% *des individus sont vaccinés contre la fièvre jaune,* 60% *sont vaccinés contre la diphtérie et* 30% *sont vaccinés contre les deux maladies.*
Quelle est la probabilité, pour un individu choisi au hasard, de n'être vacciné contre aucune de ces maladies ?
Pour traiter cet exemple, nous allons d'abord nommer les événements et on a
F :" L'individu est vacciné contre la fièvre jaune".
D :" L'individu est vacciné contre la diphtérie".
Nous avons aussi

$$\mathbf{P}(F) = \frac{45}{100} = 0.45,\ \mathbf{P}(D) = \frac{60}{100} = 0.6 \text{ et } \mathbf{P}(F \cap D) = \frac{30}{100} = 0,3.$$

La question est de calculer $\mathbf{P}(\bar{F} \cap \bar{D})$ *donc*

$$\mathbf{P}(\bar{F} \cap \bar{D}) = \mathbf{P}(\overline{F \cup D}) = 1 - \mathbf{P}(F \cup D) = 1 - [\mathbf{P}(F) + \mathbf{P}(D) - \mathbf{P}(F \cap D)]$$

et on a

$$\mathbf{P}(\bar{F} \cap \bar{D}) = 1 - [0.45 + 0.60 - .30] = 0.25.$$

La seconde propriété que nous allons voir est utiliser dans le cas où nos événements sont compatibles et en plus nous connaissons la probabilité de l'un des événements ainsi que la probabilité qu'ils se produisent en même

temps. Cette probabilité est dite **probabilité conditionnelle**. Elle est donnée par

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}, \tag{1.4}$$

pour A et B deux événements de $\mathcal{A}$ avec $\mathbf{P}(B) \neq 0$. De la formule (1.4), on peut extraire deux autres formules, mais la plus importante est donnée par

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B).\mathbf{P}(B). \tag{1.5}$$

Elle représente la formule de la **probabilité composé.**

Exemple 2 *Un service hospitalier reçoit des malades atteints soit de la maladie A, soit de la maladie B. Les proportions sont de* 40% *pour A et* 60% *pour B. Parmi les malades de A, il y a* 50% *de fumeurs et parmi ceux qui sont atteints de B, il y a* 60% *de fumeurs. Quel est le pourcentage de fumeurs dans l'ensemble des malades ? Quel est le pourcentage des non fumeurs ?*
Soient A :"Les malades atteints de la maladie A", B :"Les malades atteints de la maladie B" et F :"Les malades fumeurs".
Nous avons $\mathbf{P}(A) = 0.4$, $\mathbf{P}(B) = 0.6$ *et* $\mathbf{P}(F|A) = 0.5$, $\mathbf{P}(F|B) = 0.6$, *nous avons aussi*

$$\mathbf{P}(F) = \mathbf{P}(F \cap A) + \mathbf{P}(F \cap B) = \mathbf{P}(A)\mathbf{P}(F|A) + \mathbf{P}(B)\mathbf{P}(F|B) = 0.56,$$

et

$$\mathbf{P}(\bar{F}) = 1 - \mathbf{P}(F) = 1 - 0.56 = 0.44.$$

Dans la partie qui va suivre, nous allons essayer d'expliciter le cas de figure où les événements que nous étudions soient non dépendants l'un de l'autre.

1.3.4 Les événements indépendants

Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace **probabilisé**, soient A et$B$ deux événements. On dit que A et B sont **indépendants** si et seulement si

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \times \mathbf{P}(B). \tag{1.6}$$

Cette équation mène à des propriétés assez intéressantes utilisées dans le calcul des probabilités.

Propriétés

Soient $A, B \in \mathcal{A}$. Commençons par donner les propriétés suivantes

1. A et B indépendants $\Longrightarrow$ A et $\bar{B}$ indépendants.
2. A et B indépendants $\Longrightarrow$ $\bar{A}$ et B indépendants.
3. A et B indépendants$\Longrightarrow$ $\bar{A}$ et $\bar{B}$ indépendants.
4. A et B indépendants $\Longrightarrow \mathbf{P}(A|B) = \mathbf{P}(A)$ avec $\mathbf{P}(B) \neq 0$.

Pour avoir une idée de démonstration, nous allons voir la première propriété.

$$\mathbf{P}(A) = \mathbf{P}(A \cap \Omega) = \mathbf{P}(A \cap (B \cup \bar{B})) = \mathbf{P}(A \cap B) + \mathbf{P}(A \cap \bar{B}),$$

du fait que $(A \cap B)$ et $(A \cap \bar{B})$ sont deux événements incompatibles. Et nous avons que A et B sont indépendants, donc

$$\mathbf{P}(A) = [\mathbf{P}(A) \times \mathbf{P}(B)] + \mathbf{P}(A \cap \bar{B})$$

$$\Longrightarrow \mathbf{P}(A \cap \bar{B}) = \mathbf{P}(A) - [\mathbf{P}(A) \times \mathbf{P}(B)] = \mathbf{P}(A)\left[1 - \mathbf{P}(B)\right],$$

$$\Longrightarrow \mathbf{P}(A \cap \bar{B}) = \mathbf{P}(A)\mathbf{P}(\bar{B}).$$

Pour la quatrième affirmation nous allons utiliser la définition de la probabilité conditionnelle ainsi que celle de l'indépendance et on obtient

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)} = \frac{\mathbf{P}(A) \times \mathbf{P}(B)}{\mathbf{P}(B)} = \mathbf{P}(A).$$

Passons maintenant à la généralisation de cette formule à trois événements. Soient A, B et D des événements indépendants dans leur ensemble, ce qui veut dire qu'ils sont indépendants deux à deux. On écrit

$$\mathbf{P}(A \cap B \cap D) = \mathbf{P}(A) \times \mathbf{P}(B) \times \mathbf{P}(D).$$

Généralisation

On dit que les événements $(A_i)_{i=1,\ldots,n}$ sont indépendantes 2 à 2 si est seulement si $\forall i \neq j : \mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$.
On dit que les événements $(A_i)_{i=1,\ldots,n}$ sont mutuellement indépendants si est seulement si $\forall k = 2, ..., n$: et tout choix d'indices on a $\mathbf{P}(\cap_{j=1}^{k} A_j) = \prod_{j=1}^{k} \mathbf{P}(A_j)$.

Proposition 1 *L'indépendance mutuelle entraîne leur indépendance 2 à 2.*

La partie qui va suivre représente une formule très importante utilisée surtout dans le cas d'un système d'événements qui vérifie des conditions précises.

Formule de Bayes

Soient $(A_i)_{i=1,\dots,n}$ une suite d'événements de $\mathcal{A}$ avec $\mathbf{P}(A_i) \neq 0$. On dit de cette série que c'est un système complet de Ω si les A_i sont incompatibles deux à deux et en plus on a $\Omega = \bigcup_{i=1}^{n} A_i$. Soit $B \in \mathcal{A}$, alors

$$\mathbf{P}(B) = \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}(B|A_i). \tag{1.7}$$

Pour démontrer cette écriture nous allons utiliser la définition de la probabilité conditionnelle ainsi que celle de l'indépendance. On a l'écriture suivante

$$\begin{aligned}
\mathbf{P}(B) =& \mathbf{P}\left(B \cap \Omega\right) \\
=& \mathbf{P}\left(B \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) \\
=& \mathbf{P}\left(\bigcup_{i=1}^{n} (B \cap A_i)\right) \\
=& \sum_{i=1}^{n} \mathbf{P}\left(B \cap A_i\right) \\
=& \sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}(B|A_i).
\end{aligned}$$

Si $\mathbf{P}(B) > 0$ on peut aussi écrire

$$\mathbf{P}(A_i|B) = \frac{\mathbf{P}(A_i) \times \mathbf{P}(B|A_i)}{\sum_{i=1}^{n} \mathbf{P}(A_i) \times \mathbf{P}(B|A_i)}, \forall i = 1, \dots, n. \tag{1.8}$$

Remarque 3 *Dans le cas particulier où le système complet est $A, \bar{A}$ on obtient*

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A) \times \mathbf{P}(B|A)}{\mathbf{P}(A) \times \mathbf{P}(B|A) + \mathbf{P}(\bar{A}) \times \mathbf{P}(B|\bar{A})}. \tag{1.9}$$

Il faut savoir aussi utiliser un schéma pour mieux comprendre un exemple sur la formule de Bayes.

Exemple 3 *Un laboratoire a mis au point un alcootest. On sait que* 2% *des personnes contrôlées par la police sont réellement en état d'ébriété. Les premières essais ont conduit aux résultats suivants*

- *Lorsqu'une personne est réellement en état d'ébriété, 95 fois sur 100 l'alcootest se révèle positif.*
- *Lorsqu'une personne n'est en état d'ébriété, 96 fois sur 100 l'alcootest se révèle négatif.*

Quelle est la probabilité pour qu'une personne n'est pas en état d'ébriété lorsque l'alcootest est positif ?
Quelle est la probabilité pour qu'une personne soit réellement en état d'ébriété lorsque l'alcootest est positif ?
Notons A : " La personne contrôlée est en état d'ébriété" et $\bar{A}$:" La personne contrôlée n'est pas en état d'ébriété."
Soit B : " L'alcootest est positif" et $\bar{B}$: " l'alcootest est négatif", on a $\mathbf{P}(A) = 0.02, \mathbf{P}(B|A) = 0.95; \mathbf{P}(\bar{B}|\bar{A}) = 0.96$.
Nous voulons calculer en premier lieu $\mathbf{P}(B|\bar{A})$. *On a*

$$\mathbf{P}(\Omega|\bar{A}) = \frac{\mathbf{P}(\Omega \cap \bar{A})}{\mathbf{P}(\bar{A})} = \frac{\mathbf{P}(\bar{A})}{\mathbf{P}(\bar{A})} = 1.$$

Et

$$\begin{aligned}\mathbf{P}(\Omega|\bar{A}) =& \frac{\mathbf{P}(\Omega \cap \bar{A})}{\mathbf{P}(\bar{A})} \\ =& \frac{\mathbf{P}((B \cup \bar{B}) \cap \bar{A})}{\mathbf{P}(\bar{A})} \\ =& \frac{\mathbf{P}(B \cap \bar{A}) + \mathbf{P}(\bar{B} \cap \bar{A})}{\mathbf{P}(\bar{A})} \\ =& \mathbf{P}(B|\bar{A}) + \mathbf{P}(\bar{B}|\bar{A}).\end{aligned}$$

Donc

$$\mathbf{P}(B|\bar{A}) + \mathbf{P}(\bar{B}|\bar{A}) = 1. \tag{1.10}$$

Cette relation démontre que la somme des probabilités conditionnelle par rapport à un ensemble est égale à un.

Enfin, on obtient que $\mathbf{P}(B|\bar{A}) = 1 - \mathbf{P}(\bar{B}|\bar{A}) = 1 - 0.96 = 0.04$.
Et nous voulons calculer $\mathbf{P}(A|B)$. *D'après la formule (1.9), on a*

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A) \times \mathbf{P}(B|A)}{\mathbf{P}(A) \times \mathbf{P}(B|A) + \mathbf{P}(\bar{A}) \times \mathbf{P}(B|\bar{A})} = 0.3265.$$

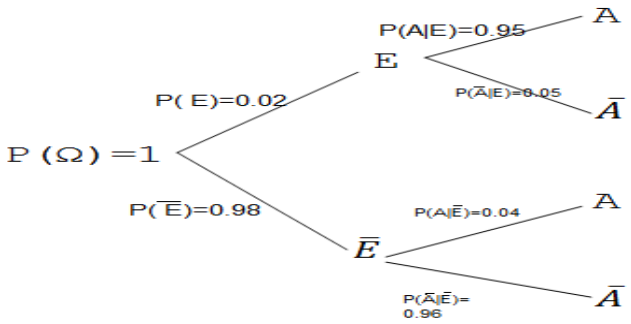


Figure 1.1 – Schéma de la formule de Bayes.

1.3.5 Calcul de probabilité non uniforme

Lorsque nous somme dans le cas non équiprobable, nous utilisons les mêmes propriétés que le cas uniforme, mais pour calculer le cardinal d'un événement, nous utilisons l'analyse combinatoire et l'exemple qui suit nous donne un aperçut de la méthode.

Exemple 4 *18 personnes se sont présentées à une collecte de sang. Parmi celles-ci, on a noté 11 personnes du groupe O, 4 personnes du groupes A, 2 personnes du groupe B et 1 personne du groupe AB.*
A l'issue de la collecte, on prélève au hasard 3 flacons parmi les 18 flacons obtenus. Calculer la probabilité des événements suivants :
a)

- *Les trois flacons appartiennent au groupe O.*
- *Les trois flacons appartiennent au groupe A.*
- *Les trois flacons appartiennent au groupe B.*

— *Les trois flacons appartiennent au groupe AB.*

b) Les trois flacons appartiennent au même groupe.

c) Parmi les 3 flacons, il y' a au moins 1 flacon du groupe A.

d) Les trois flacons appartiennent à 3 groupes différents.

Il faut savoir décortiquer ce genre d'exemple pour pouvoir répondre aux questions. La première chose à faire est de désigner Ω *ou bien calculer son cardinal sans la désigner. Ici* $\Omega = \{(O, A, B), \ldots\}$*, donc elle représente l'ensemble des triplets du genre* (O, A, B) *et on voit bien que nous ne pouvons l'écrire complètement. Il ne reste qu'a calculer son cardinal en utilisant l'analyse combinatoire.*

Remarquer que pour choisir les événements du genre (O, A, B)*, nous n'avons pas précisé l'emplacement de chacun. Donc nous n'avons pas d'ordre, ce qui revient à utiliser une combinaison et ainsi, on obtient que*

$$card(\Omega) = C_{18}^{3} = 816.$$

La deuxième étape est de nommer les événements que nous voulons utiliser. Soient

— *C :"Les trois flacons appartiennent au groupe O".*

— *D :"Les trois flacons appartiennent au groupe A".*

— *E :" Les trois flacons appartiennent au groupe B".*

— *F :"Les trois flacons appartiennent au groupe AB".*

a) Reste à calculer leurs probabilités. On a

$$\mathbf{P}(C) = \frac{card(C)}{card(\Omega)} = \frac{C_1^3 1}{C_{18}^3} = \frac{165}{816}.$$

$$\mathbf{P}(D) = \frac{card(D)}{card(\Omega)} = \frac{C_4^3}{C_{18}^3} = \frac{4}{816}.$$

$$\mathbf{P}(E) = \frac{card(E)}{card(\Omega)} = \frac{C_2^3}{C_{18}^3} = \frac{0}{816} = 0.$$

$$\mathbf{P}(F) = \frac{card(F)}{card(\Omega)} = \frac{C_1^3}{C_{18}^3} = \frac{0}{816} = 0.$$

b) Les trois flacons appartiennent au même groupe,veut dire que les trois flacons proviennent (du groupe $O \bigvee$ *du groupe A). Ils ne peuvent provenir du groupe B et AB du fait qu'il y a moins de trois flacons. Donc nous allons*

calculer $\mathbf{P}(C \cup E)$. *On obtient* $\mathbf{P}(C \cup E) = \mathbf{P}(C) + \mathbf{P}(E)$ *et cela du fait que* C *et* E *sont deux événements incompatibles* $(C \cap E) = \emptyset$. *Ainsi*

$$\mathbf{P}(C \cup E) = \mathbf{P}(C) + \mathbf{P}(E) = \frac{165}{816} + \frac{4}{816} = \frac{169}{816}.$$

c) Parmi les 3 flacons, il y' a au moins 1 flacon du groupe A, veut dire qu'il y a [ou bien (R_1 *:1 flacon du groupe A)* $\bigvee$ (R_1 : *2 flacons du groupe A)* $\bigvee$ (R_1 : *3 flacons du groupe A)] et* R_1, R_2 *et* R_3 *sont des événements incompatibles donc nous allons calculer la somme des probabilités*

$$\mathbf{P}(R_1 \cup R_2 \cup R_3) = \mathbf{P}(R_1) + \mathbf{P}(R_2) + \mathbf{P}(R_3).$$

Avec

$$\mathbf{P}(R_1) = \frac{card(R_1)}{card(\Omega)} = \frac{C_4^1 \times C_{(18-4)}^2}{C_{18}^3} = \frac{364}{816}.$$

$$\mathbf{P}(R_2) = \frac{card(R_2)}{card(\Omega)} = \frac{C_4^2 \times C_{(18-4)}^1}{C_{18}^3} = \frac{6}{816}.$$

$$\mathbf{P}(R_3) = \frac{card(R_3)}{card(\Omega)} = \mathbf{P}(E) = \frac{4}{816}.$$

Au final, on obtient

$$\mathbf{P}(R_1 \cup R_2 \cup R_3) = \frac{364}{816} + \frac{6}{816} + \frac{4}{816}.$$

d) H :"Les trois flacons appartiennent à 3 groupes différents", équivaut $[O, A, B] \bigvee [O, A, AB] \bigvee [O, B, AB] \bigvee [A, B, AB]$, *donc*

$$\mathbf{P}(H) = \frac{C_4^1 \times C_1^1 1 \times C_2^1}{816} + \frac{C_4^1 \times C_1^1 1 \times C_1^1}{816} + \frac{C_1^1 \times C_1^1 1 \times C_2^1}{816} + \frac{C_4^1 \times C_2^1 \times C_1^1}{816}$$

$$= \frac{162}{816}.$$

Les chapitres suivants représentent une ouverture vers les variables aléatoires usuelles dans le cas discret et continu.

1.4 Exercices

Exercice 1 *Soient $(\Omega, \mathcal{A})$ un espace probabilisable, A, B deux événements de $\mathcal{A}$. Donner pour chaque assertion son écriture en symboles ($\cup$, $\cap$, complémentaire).*

1. *A se réalise.*
2. *L'événement contraire de A se réalise.*
3. *A et B se réalisent.*
4. *A et B ne se réalisent pas.*
5. *A et le contraire de B se réalisent.*
6. *un seul événement des deux se réalise.*
7. *Au moins l'un des deux se réalise.*

Exercice 2 *Soient $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé, A, B et $(A_n)_{n\in\mathbb{N}}$ des événements de $\mathcal{A}$. Montrer les assertions suivantes.*

1. $\mathbf{P}(\emptyset) = 0$.
2. $\mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A)$.
3. $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.
4. *Si $A \subset B$ alors* $\mathbf{P}(A) \leq \mathbf{P}(B)$.
5. $\forall n \in \mathbb{N},\ \mathbf{P}\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) \leq \sum_{n\in\mathbb{N}} \mathbf{P}(A_n)$.
6. *Si $(A_n)_{n\in\mathbb{N}}$ croissante alors* $\mathbf{P}\left(\underset{n\in\mathbb{N}}{\cup} A_n\right) = \lim_{n\to\infty} \mathbf{P}(A_n)$.
7. *Si $(A_n)_{n\in\mathbb{N}}$ décroissante alors* $\mathbf{P}\left(\underset{n\in\mathbb{N}}{\cap} A_n\right) = \lim_{n\to\infty} \mathbf{P}(A_n)$.

Exercice 3 *Dans une population, 45% des individus sont vaccinés contre la fièvre jaune, 60% sont vaccinés contre la diphtérie et 30% sont vaccinés contre les deux maladies. Quelle est la probabilité, pour un individu choisi au hasard, de n'être vacciné contre aucune de ces maladies ?*

Exercice 4 *Soient $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé, A, B deux événements de $\mathcal{A}$. Montrer les équivalences suivantes.*

1. *A et B indépendants.*
2. *A et $\bar{B}$ indépendants.*
3. *$\bar{A}$ et B indépendants.*
4. *$\bar{A}$ et $\bar{B}$ indépendants.*

Exercice 5 *Le tiers d'une population a été vacciné contre une maladie contagieuse. au cours d'une épidémie, on a constaté que 5% des malades avaient été vaccinés et que 8% des personnes vaccinées ont été malades.*

1. *Quelle était la probabilité pour un individu de la population de tomber malade lors de l'épidémie ?*
2. *Quelle était la probabilité de tomber malade pour un individu non vacciné ?*

Exercice 6 *Au retour de la plage, dans un groupe de 20 vacanciers, 5 ont eu des coups de soleil, 8 ont été piqués par les moustiques et 10 sont saufs.*

1. *Quelle est la probabilité qu'un de ces touristes tiré au hasard souffre à la fois d'un coup de soleil et soit piqué par les moustiques.*
2. *Sachant que ce touriste a été piqué par les moustiques, quelle est la probabilité qu'il ait échappé aux coups de soleil ?*
3. *Notons M :"être piqué par les moustiques" et S :"avoir un coup de soleil". Les événements M et S sont-ils indépendants ? Comparer $\mathbf{P}(M|S)$ et $\mathbf{P}(M|\overline{S})$.*

Exercice 7 *Soient $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé, A, B et C des événements de $\mathcal{A}$ tel que $\mathbf{P}(A \cap C) \neq 0$.*

1. *Montrer que $\mathbf{P}(A \cap C|B) = \mathbf{P}(A|B \cap C)\mathbf{P}(C|B)$.*
2. *Supposons maintenant que les événements A, B et C sont mutuellement indépendants tels que $\mathbf{P}(A) = 0.1$, $\mathbf{P}(B) = 0.3$ et $\mathbf{P}(C) = 0.2$. Calculer :$\mathbf{P}(\bar{A} \cup (B \cap C))$ et $\mathbf{P}(A \cup \bar{B} \cup C)$.*

Exercice 8 *Un service hospitalier reçoit des malades atteints soit de la maladie A, soit de la maladie B. Les proportions sont de 40% pour A et 60% pour B. Parmi les malades de A, il y a 50% de fumeurs et parmi ceux*

qui sont atteints de B, il y a 60% *de fumeurs. Quel est le pourcentage de fumeurs dans l'ensemble des malades ? Quel est le pourcentage des non fumeurs ?*

Exercice 9 *Un laboratoire a mis au point un alcootest. On sait que* 2% *des personnes contrôlées par la police sont réellement en état d'ébriété. Les premières essais ont conduit aux résultats suivants*

- *Lorsqu'une personne est réellement en état d'ébriété, 95 fois sur 100 l'alcootest se révèle positif.*
- *Lorsqu'une personne n'est en état d'ébriété, 96 fois sur 100 l'alcootest se révèle négatif.*

Quelle est la probabilité pour qu'une personne n'est pas en état d'ébriété lorsque l'alcootest est positif ?
Quelle est la probabilité pour qu'une personne soit réellement en état d'ébriété lorsque l'alcootest est positif ?

2 Les variables aléatoires discrètes et leur lois usuelles

La définition moderne d'une v.a. ne peut être exposée rigoureusement sans faire appel à la théorie de la mesure et de l'intégration au sens de Lebesgue. Donc dans ce chapitre, on définit d'une manière moins rigoureuse une variable aléatoire. Aussi, on expose la fonction de répartition et les moments d'ordre un et deux. Quelques lois discrètes, avec leurs paramètres sont introduites.

2.1 Définition générale d'une variable aléatoire

Pour f défini de E dans F avec $B \subset F$ on appel image réciproque de B par f l'ensemble donné par $f^{-1}(B) = \{x \in E|\ f(x) \in B\}$.

Définition 1 *Soient $(E; \mathcal{A})$ et $(F; \mathcal{F})$ deux espaces mesurables, on dira que la fonction f définie de $(E; \mathcal{A})$ dans $(F; \mathcal{F})$ est mesurable si pour tout $B \in \mathcal{F}$ alors $f^{-1}(B)$ appartient à $\mathcal{A}$.*

Définition 2 *Soit $(\Omega, \mathcal{A}, \mathbf{P})$ un espace probabilisé (e.p.) et $X : \Omega \to \mathbb{R}$. On dit que X est mesurable (voir la définition1) si pour tout $x \in \mathbb{R}$ l'ensemble $X^{-1}(]-\infty,\ x]) \in \mathcal{A}$. On nome X variable aléatoire réelle noté v.a.r. où $\mathcal{B}(\mathbb{R})$ est dite tribu borélienne sur $\mathbb{R}$ (ou tribu des boréliens de $\mathbb{R}$) qui est engendrée par les intervalles ouverts $(\forall x,]-\infty,\ x])$ de $\mathbb{R}$.*

Remarque 4

1. $X^{-1}(]-\infty,\ x]) = \{\omega \in \Omega|\ X(\omega) \in]-\infty,\ x]\} = \{\omega \in \Omega|\ X(\omega) \leq x\}$.

2. *On utilise l'écriture suivante* $\forall A \in \mathcal{B}(\mathbb{R});\ X^{-1}(A) = \{X \in A\}$. *C-à-d qu'on considère* ω *comme indice caché.*

3. *Soient* X *et* Y *deux v. a. r.* $\Longrightarrow X+Y,\ X-Y,\ X.Y,\ \frac{X}{Y}$ *sont des variables aléatoires.*

4. *Comme* $X^{-1}(]-\infty,\ x]) \in \mathcal{A} \Longrightarrow \mathbf{P}(X^{-1}(]-\infty,\ x])$ *existe. C-à-d, on peut calculer sa probabilité.*

5. *Si* $X(\Omega) =$ *un ensemble fini ou dénombrable, on dit que* X *est une v.a.r. discrète.*

6. *Si* $X(\Omega) = \mathbb{R}$ *ou intervalle de* $\mathbb{R}$*, on dit que* X *est une v.a.r continue.*

2.1.1 La loi de probabilité

Définition 3 *Soient* $(\Omega, \mathcal{A}, \mathbf{P})$ *e.p., X une v.a.r. Soit* $\mathbf{P}_X$ *définie sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *par*

$$\forall A \in \mathcal{B}(\mathbb{R});\ \mathbf{P}_X(A) = \mathbf{P}[\omega \in \Omega : X(\omega) \in A] = \mathbf{P}[X^{-1}(A)].$$

$\mathbf{P}_X$ *est dite loi de probabilité de* X.

Remarque 5

1. *Ici aussi on utilise les notations*

$$\mathbf{P}_X(A) = \mathbf{P}[\omega \in \Omega : X(\omega) \in A] = \mathbf{P}[X \in A]$$

et $\mathbf{P}_X\left(X^{-1}(]-\infty,\ x])\right) = \mathbf{P}(X \leq x)$.

2. *Aussi* $\mathbf{P}_X(\{x\}) = \mathbf{P}[\omega \in \Omega : X(\omega) \in \{x\}] = \mathbf{P}[X = x]$.

Proposition 2 *L'application* $\mathbf{P}_X$ *est une probabilité sur* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

2.2 Variable aléatoire discrète

Dans la suite nous allons définir la loi, la fonction de répartition et les moments d'une variable aléatoire discrète.

2.2.1 Loi d'une v.a. discrète

Si X une v.a. discrète de valeurs $\{x_1, \ldots, x_n, \ldots\}$, on dit qu'elle est de loi de probabilité $(p_i = \mathbf{P}(X = x_i))_{i=1,\ldots,n, \ldots}$. On a aussi,

$$\sum_{i \geq 1} \mathbf{P}(X = x_i) = 1.$$

L'exemple suivant montre l'utilisation du passage de l'ensemble des événements vers l'ensemble réel à l'aide du codage X.

Exemple 5 *Soit l'expérience aléatoire :" jet de deux dès non pipés". On obtient* $\Omega = \{(1,1); (1,2); \ldots; (6,6)\}$.

1. *On considère* X *:"somme des faces"*$\Leftrightarrow X = \{2, 3, \ldots, 12\}$. *On calcule la probabilité de chaque valeur de* X.
   $\mathbf{P}(X = 2) = \mathbf{P}[(1,1)] = \frac{1}{36}$;
   $\mathbf{P}(X = 3) = \mathbf{P}[(1,2); (2,1)] = \frac{2}{36}$;
   $\mathbf{P}(X = 4) = \mathbf{P}[(2,2); (3,1); (1,3)] = \frac{3}{36}$;
   $\mathbf{P}(X = 5) = \mathbf{P}[(1,4); (4,1); (3,2); (2,3)] = \frac{4}{36}$;
   $\mathbf{P}(X = 6) = \mathbf{P}[(1,5); (5,1); (3,3); (4,2), (2,4)] = \frac{5}{36}$;
   $\mathbf{P}(X = 7) = \mathbf{P}[(3,4); (4,3); (5,2); (2,5); (1,6); (6,1)] = \frac{6}{36}$;
   $\mathbf{P}(X = 8) = \mathbf{P}[(2,6); (6,2); (3,5); (5,3); (4,4)] = \frac{5}{36}$;
   $\mathbf{P}(X = 9) = \mathbf{P}[(3,6); (6,3); (4,5); (5,4);] = \frac{4}{36}$;
   $\mathbf{P}(X = 10) = \mathbf{P}[(5,5); (4,6); (6,4)] = \frac{3}{36}$;
   $\mathbf{P}(X = 11) = \mathbf{P}[(5,6); (6,5)] = \frac{2}{36}$;
   $\mathbf{P}(X = 12) = \mathbf{P}[(6,6)] = \frac{1}{36}$.
2. *On considère* Y *:"maximum des deux faces"*$\Leftrightarrow Y = \{1, 2, \ldots, 6\}$.
   $\mathbf{P}(Y = 1) = \mathbf{P}[(1,1)] = \frac{1}{36}$;
   $\mathbf{P}(Y = 2) = \mathbf{P}[(2,2); (1,2); (2,1)] = \frac{3}{36}$;
   $\mathbf{P}(Y = 3) = \mathbf{P}[(1,3); (3,1); (2,3); (3,2); (3,3)] = \frac{5}{36}$;
   $\mathbf{P}(Y = 4) = \mathbf{P}[(1,4); (4,1); (2,4); (4,2); (3,4); (4,3); (4,4)] = \frac{7}{36}$;
   $\mathbf{P}(Y = 5) = \mathbf{P}[(1,5); (5,1)(2,5), (5,2); (3,5); (5,3); (4,5); (5,4);$
   $(5,5)] = \frac{9}{36}$;
   $\mathbf{P}(Y = 6) = \mathbf{P}[(1,6); (6,1); (2,6); (6,2); (3,6); (6,3); (4,6); (6,4); (5,6);$
   $(6,5); (6,6)] = \frac{11}{36}$.

2.3 Fonction de répartition d'une v.a. discrète

On appelle fonction de répartition de X la fonction noté F_X définie de $\mathbb{R}$ dans $[0,\ 1]$ par

$$F_X(x) = \mathbf{P}(]-\infty,\ x]) = \mathbf{P}(X \leq x);\ \forall x \in \mathbb{R}. \tag{2.1}$$

F_X vérifie les propriétés suivantes

1. $\forall x \in \mathbb{R},\ 0 \leq F_X(x) \leq 1$.
2. F_X est une fonction croissante.
3. $\forall a, b \in \mathbb{R} : \mathbf{P}\left[X \in]a,\ b]\right] = F_X(b) - F_X(a)$.
4. F_X est une fonction continue à droite et admet une limite à gauche en tout point $x \in \mathbb{R}$.
   $\Leftrightarrow \lim\limits_{\substack{>\\ h\to 0}} F_X(x+h) = \lim\limits_{\substack{>\\ h\to 0}} \mathbf{P}(X \leq x+h) = \mathbf{P}(X \leq x) = F_X(x)$,
   $\Leftrightarrow \lim\limits_{\substack{<\\ h\to 0}} F_X(x+h) = \lim\limits_{\substack{<\\ h\to 0}} \mathbf{P}(X \leq x+h) = \mathbf{P}(X < x) = F_X(x-)$.

Remarque 6 *1. Toute fonction qui vérifie les propriétés précédentes est une fonction de répartition.*

2. Deux fonctions de répartitions égales caractérisent la même loi de probabilité.

En utilisant la définition de la fonction de répartition donnée par la formule (2.1), on a $\forall x \in \mathbb{R}, F_X(x) = \mathbf{P}(X \leq x) = \sum_{x_i \leq x} p_i$. On obtient une fonction étagée continue à droite avec limite à gauche (càdlàg). Représentée dans la figure (2.1).

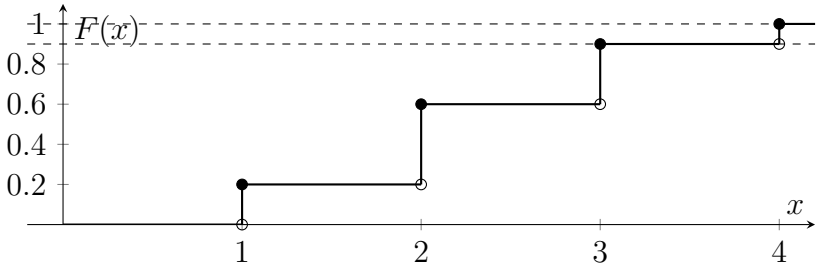


Figure 2.1 – Fonction de répartition (diagramme en escalier)

2.4 Les moments

L'espérance d'une v.a. joue le même rôle que la moyenne d'une v. statistique (v.s.). C'est-à-dire que la formule de l'espérance s'inspire de celle de la moyenne. Soient X une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ de loi $(x_i, p_i)_{i=1,\dots n}$, on a

$$\bar{x} = \sum_{i=1}^{n} f_i x_i \text{ dans le cas descriptive} \Leftrightarrow \mathbf{E}(X) = \sum_{i=1}^{n} p_i x_i \text{ dans le cas aléatoire.}$$

Ici f_i représente la fréquence d'observé x_i et p_i la probabilité de l'observé. Calculer les moments d'ordre p d'une variable équivaut le calcul des moyennes (noté $\bar{x}$ dans le cas descriptive $\Leftrightarrow$ noté $\mathbf{E}(X)$ dans le cas aléatoire) de la puissance p (c-à-d X^p) de cette v.a. mais il faut qu'elle soit centrée (c-à-d $\mathbf{E}\{(X - \mathbf{E}(X))^p\}$).
Le moment d'ordre 2 représente la variance de la v.a. qui s'écrit

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

Ce qui implique que

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

Remarquer qu'elle s'écrit de la même manière que dans le cas d'une v. s.

Pour calculer les moments d'ordre p d'une v.a. X il faut calculer $\mathbf{E}(X^p)$. D'une manière générale, soit $Z = g(X)$ une v.a. fonction de la v. a. X. Pour calculer $\mathbf{E}(Z)$ on peut d'abord déterminer sa loi à partir de celle de X et ensuite utiliser la définition de l'espérance mathématique. Mais, il est possible de montrer que l'on peut directement calculer $\mathbf{E}(Z)$ sur la loi de X en utilisant le théorème suivant.

Théorème 1 *Soit X une v. a. discrète à valeurs $x_1; x_2; \dots$ et g une application mesurable de $(\mathbb{R}; \mathcal{B}_\mathbb{R})$ dans $(\mathbb{R}; \mathcal{B}_\mathbb{R})$, alors la v. a $Z = g(X)$ admet une espérance si et seulement si $\sum_{i\geq 1} |g(x_i)|\mathbf{P}(X = x_i) < \infty$ et alors*

$$\mathbf{E}(Z) = \mathbf{E}(g(X)) = \sum_{i\geq 1} g(x_i)\mathbf{P}(X = x_i).$$

Ce qui implique que $\mathbf{E}(X^p) = \sum_{i\geq 1} x_i^p \mathbf{P}(X = x_i),\ \forall p \geq 1$. Une formule très utile pour calculer la variance d'une v.a.

De la linéarité de la fonction somme, on a la linéarité de la fonction espérance.

Propriété 1 *1. Soient g et h deux fonctions de $\mathbb{R}$ et $a, b \in \mathbb{R}$, on a*

$$\mathbf{E}(ag(X) + bh(X)) = a\mathbf{E}(g(X)) + b\mathbf{E}(h(X)).$$

2. Aussi, pour $a, b \in \mathbb{R}$, on a

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b.$$

3. Pour $a \in \mathbb{R}$, on a $\mathbf{E}(a) = a$.

4. Le moment centré d'ordre $k \in \mathbb{N}^*$ de la v. a. X est donné par

$$\mathbf{E}((X - \mathbf{E}(X))^k).$$

Donc, on a $\mathbf{E}((X-\mathbf{E}(X))^1) = 0$ et $\mathbf{E}((X-\mathbf{E}(X))^2) = \mathbf{E}(X^2)-\mathbf{E}^2(X)$. D'où que

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

5. Pour $a \in \mathbb{R}$, on a : $\mathbf{Var}(a) = \mathbf{E}(a^2) - (\mathbf{E}^2(a)) = a^2 - a^2 = 0$.

6. Pour $a \in \mathbb{R}$, on a

$$\mathbf{Var}(aX) = \mathbf{E}(a^2X^2)-(\mathbf{E}(aX))^2 = a^2\mathbf{E}(X^2)-a^2\mathbf{E}^2(X) = a^2\mathbf{Var}(X).$$

7. L'écart-type de X est donné par $\sigma_X = \sqrt{\mathbf{Var}(X)}$.

Il existe trois façons pour caractériser la loi d'une variable aléatoire discrète (voir exercice (17)). La première est la définition (2.1.1) d'une v. a. r., la deuxième est la fonction de répartition et la troisième est la fonction de masse ou poids $\mathbf{P}[X = k]$. La question qui se pose est, es-qu'il y a une manière générale pour le faire ?

2.5 Fonctions génératrices des moments

Soient X une v. a. r. et $\{\mathbf{E}(g(X)), g \in \mathcal{F}\}$, avec $\mathcal{F}$ est une classe de fonction assez riche. Par exemple

1. $g = \mathbf{1}_{]-\infty,\ x]},\ x \in \mathbb{R}$.
2. g une fonction continue bornée $g = \lim_{n\to\infty} \mathbf{1}_{A_n}$.
3. g fonction positive définie par $g = \sum_{i=1}^{n} \mathbf{1}_{A_n}$.
4. g fonction trigonométrique $g(x) = e^{itx},\ t \in \mathbb{R}$ dite transformé de Fourier.
5. $g(x) = s^x,\ s \in [0,\ 1]$.
6. $g(x) = e^{tx},\ t \geq 0$ dite transformé de Laplace pour $x \geq 0$.

Dans le cas discret à valeurs dans $\mathbb{N}$ on utilise la fonction génératrice.

Définition 4 — *Si* $\mathbf{P}$ *est une loi de probabilité sur* $\mathbb{N}$ *alors sa fonction génératrice notée* $g_{\mathbf{P}} : [0,\ 1] \longrightarrow \mathbb{R}$ *est définie par*

$$g_{\mathbf{P}}(s) = \sum_{n=0}^{\infty} s^n \mathbf{P}(n),\ \textit{avec } s \in]-1,\ 1].$$

— *Si* X *est une v. a. discrète de loi* $\mathbf{P}$ *sur* $\mathbb{N}$ *alors* $g_X = g_{\mathbf{P}}$ *définie par*

$$g_X(s) = \sum_{n=0}^{\infty} s^n \mathbf{P}(X = n) = \mathbf{E}(X^s),\ \textit{avec } s \in]-1,\ 1].$$

Théorème 2 *Propriétés des fonctions génératrices*
Soit X *une v. a. à valeurs dans* $\mathbb{N}$. g_X *vérifie les propriétés suivantes*

1. g_X *est croissante sur* $[0,\ 1]$ *avec*

$$g_X(0) = \mathbf{P}[X = 0] \textit{ et } g_X(1) = \mathbf{P}[X = 1].$$

2. g_X *est* $\mathcal{C}^\infty(]-1,\ 1[)$ *est ses dérivées "engendrent les probabilités" :*

$$g_X^{(n)} = n!\mathbf{P}[X = n],\ \forall n \in \mathbb{N}.$$

3. *Soit* Y *une v. a. discrète ;* X *et* Y *on même loi* $\Leftrightarrow g_X = g_Y$.

4. Si $X(X-1)\dots(X-k+1)$ est intégrable ($k \in \mathbb{N}^*$) alors

$$\mathbf{E}\left(X(X-1)\dots(X-k+1)\right) = \lim_{s \nearrow 1} g_X^{(k)}(s).$$

La démonstration de ce théorème se fera comme exercice (18).

Comme il a été vu plus haut, si le nombre des valeurs de X sont fini ou dénombrable, on parle de v.a.r discrète. Il existe des lois de probabilités usuelles discrètes qui sont : la loi de Bernoulli, Binomiale, hypergéométrique, géométrique et Poisson. La suite du cours est consacré aux variables discrètes les plus utilisées ainsi que leurs caractéristiques c-à-d loi, fonction de répartition, espérance et variance.

2.6 Quelques lois discrètes usuelles

Soit X v.a.r discrète à valeurs $(x_i)_{i\in I}$ de loi de probabilité $(p_i)_{i\in I}$, tel que $p_i = \mathbf{P}[X = x_i]$ est dite fonction de masse. Pour chaque loi on donne les caractéristiques représentées par l'écriture explicite de la loi et la fonction de répartition, ici, la formule est donnée par

$$F_X(x) = \mathbf{P}[X \leq x] = \sum_{i\in I, x_i \leq x} \mathbf{P}[X = x_i] = \sum_{i\in I, x_i \leq x} p_i.$$

Cette écriture correspond à la notion de fréquence cumulée rencontrée en statistique descriptive notée F_i. Enfin, les moments d'ordre un et deux donnés par

$$\mathbf{E}(X) = \sum_{i\in I} x_i p_i,$$

$$\mathbf{Var}(X) = \mathbf{E}[(X - \mathbf{E}(X))^2] = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = \sum_{i\in I} x_i^2 p_i - (\mathbf{E}(X))^2.$$

2.6.1 Loi de Bernoulli

Si on dispose d'une expérience aléatoire avec uniquement deux événements alternatives possibles de type "succès ou échec", "vrai ou faux", "marche ou arrêt", "pile ou face",... etc, mais pas les deux simultanément.

La loi de probabilité

Un succès est représenté par l'événement $\{X = 1\}$ tandis que $\{X = 0\}$ correspond à un échec avec $X(\Omega) = \{0, 1\}$.
Puisque l'on a $\mathbf{P}[X = 0] = 1 - \mathbf{P}[X = 1]$, la loi de X ne dépend que d'un paramètre (la probabilité de succès p) ; on parle alors de la loi de Bernoulli de paramètre p noté par $X \simeq \mathcal{B}(p)$ et caractérisée par $\mathbf{P}[X = 1] = p$, $\mathbf{P}[X = 0] = 1 - p$. On peut aussi écrire

$$\mathbf{P}[X = k] = p^k(1 - p)^{1-k}, \quad \text{avec } k = 0,\ 1. \tag{2.2}$$

Cette loi est représentée par un diagramme à deux battons (voir par exemple pour $p = 0.3$ figure 2.2).

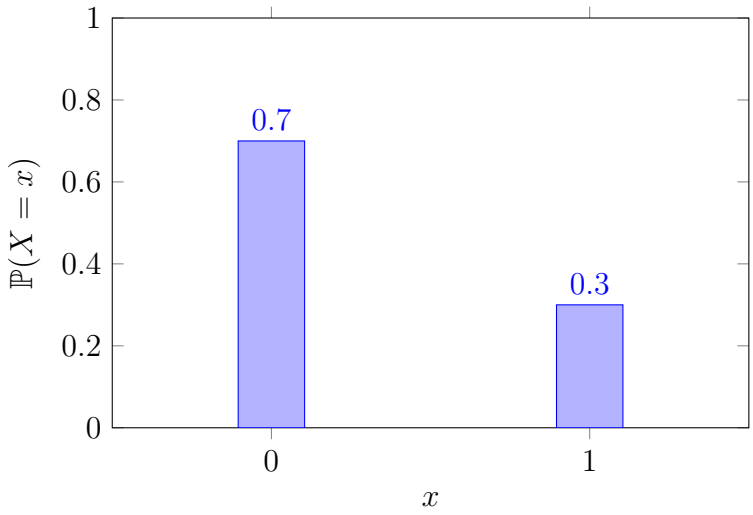


FIGURE 2.2 – Diagramme en bâtons de $X \simeq \mathcal{B}(0.3)$.

La fonction de répartition

La fonction de répartition d'une variable $X \simeq \mathcal{B}(p)$ est

$$F_X(x) = \begin{cases} 0 & \text{si } ; x < 0 \\ 1 - p & \text{si } ; 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases} .$$

Le graphe de la fonction de répartition de la loi de Bernoulli est une fonction en escalier à trois marches (voir pour le même exemple figure 2.3).

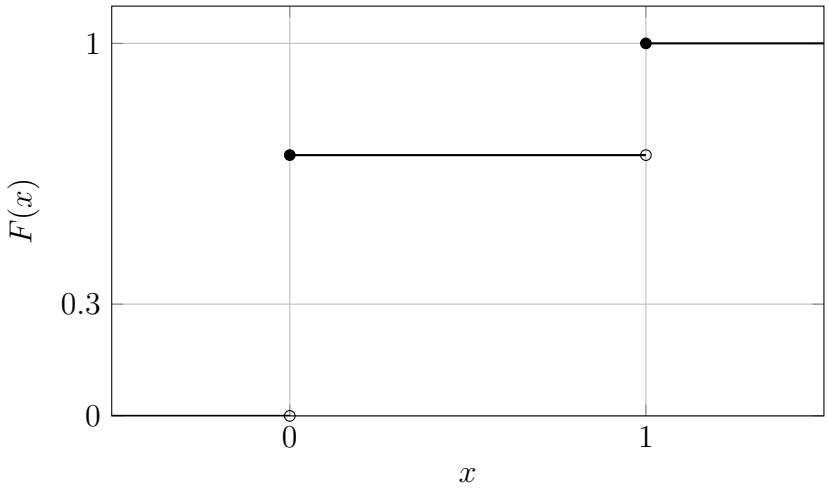


Figure 2.3 – Fonction de répartition de $X \simeq \mathcal{B}(0.3)$.

Les moments $p = 1,\ 2$

L'espérance d'une variable $X \simeq \mathcal{B}(p)$ se calcule comme suit

$$\mathbf{E}(X) = \sum_{i=0,1} x_i \mathbf{P}[X = x_i] = 0.(1-p) + 1.(p) = p.$$

Et la variance est donnée par

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

$$= \sum_{i=0,1} x_i^2 \mathbf{P}[X = x_i] - p^2 = 0.(1-p) + 1.p - p^2 = p(1-p).$$

En résumé $X \simeq \mathcal{B}(p) \Leftrightarrow$

1. $X = \{0,\ 1\}$,
2. $\mathbf{P}[X = k] = p^k(1-p)^{1-k}$, avec $k = 0,\ 1$.
3. $\mathbf{E}(X) = p, \mathbf{Var}(X) = p(1-p)$.

2.6.2 Loi Binomiale

Si on répète l'expérience de Bernoulli n fois d'une manière indépendante et on calcule le nombre de succès obtenus (c-à-d, le nombre de $X = 1$), on se retrouve avec une nouvelle v.a. notée S_n (S pour somme), dite loi Binomiale noté $S_n \simeq \mathcal{B}(n,\ p)$. Qui prend les valeurs $0,\ 1,\ \ldots,\ n$, ce qui veut dire que pour $S_n = 0$ on obtient aucun succès. Si $S_n = 1$, on obtient un succès et $(n-1)$ échecs. $\ldots$ Si $S_n = n$, on obtient un succès et aucun échec. Donc pour obtenir la loi de cette variable, on doit calculer $\mathbf{P}(S_n = k)$ avec $k = 0,\ 1,\ \ldots\ n$.

La loi de probabilité

On peut écrire S_n comme étant la somme de n v.a. indépendantes et identiquement distribuées (i. i. d) $X_i \simeq \mathcal{B}(p)$ comme suit, $S_n = \sum_{i=1}^{n} X_i$. Avec

$$X_i = \begin{cases} 0 & \text{si succès à la i}^{eme} \text{ épreuve avec une probabilité } p \\ 1 & \text{si échec à la i}^{eme} \text{ épreuve avec une probabilité } 1-p \end{cases}.$$

Cette écriture simplifiera le calcul de l'espérance et la variance. Aussi pour avoir la loi de la v. a. on doit comprendre que

- Il n'y a qu'un seul cas qui correspondant à 0 succès parmi n épreuves : (Échec, Échec, $\ldots$, Échec).
- Il n'y a qu'un seul cas qui correspondant à n succès parmi n épreuves : (Succès, Succès, $\ldots$, Succès).
- Il y a n cas qui correspondant à 1 succès parmi n épreuves : (Succès, Échec, Échec, $\ldots$, Échec) ou (Échec, Succès, Échec, $\ldots$, Échec) ou (Échec, Échec, Succès, $\ldots$, Échec)$\ldots$(Échec, Échec, Échec, $\ldots$, Succès).
- Il y a C_n^k cas qui correspond à k succès et $(n-k)$ échec parmi n épreuves.
- La probabilité de l'événement qui correspond à k succès et $(n-k)$ échec parmi n épreuve est

$$\begin{aligned} &\mathbf{P}\left[(X_1 = 1) \text{ et } (X_2 = 1)\ \ldots \text{ et } (X_k = 1) \text{ et } (X_{k+1} = 0) \ldots \text{ et } (X_n = 0)\right] \\ &\overset{\text{indépendance}}{=} \mathbf{P}(X_1 = 1)\mathbf{P}(X_2 = 1)\ldots\mathbf{P}(X_k = 1)\mathbf{P}(X_{k+1} = 0)\ldots\mathbf{P}(X_n = 0) \end{aligned}$$

$$= p^k(1-p)^{(n-k)}.$$

Donc, on obtient au final la loi de probabilité donnée par

$$\mathbf{P}[S_n = k] = C_n^k p^k (1-p)^{(n-k)} = p_k, \quad \text{avec } k = 0,\ 1, \ldots, n. \tag{2.3}$$

Cette loi est représentée par un diagramme à n battons (voir pour le même exemple figure 2.4).

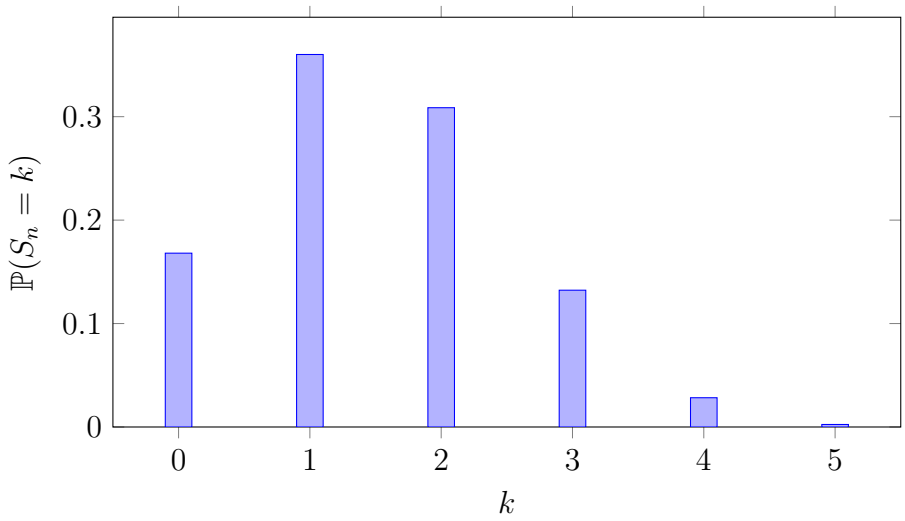


Figure 2.4 – Diagramme en bâtons de la loi binomiale $\mathcal{B}(5, 0.3)$.

La fonction de répartition

La fonction de répartition d'une variable $S_n \simeq \mathcal{B}(n,\ p)$ de valeurs $k = 0, \ldots, n$ et de loi de probabilité définie par p_k donnée par la formule (2.3) est

$$F_X(x) = \begin{cases} 0 & \text{si}\,;\ x < 0 \\ p_1 & \text{si}\,;\ 0 \leq x < 1 \\ (p_1 + p_2) & \text{si}\,;\ 1 \leq x < 2 \\ \vdots & \vdots \\ 1 & \text{si } (n-1) \leq x \leq n \end{cases}.$$

Le graphe de la fonction de répartition de la loi Binomiale est une fonction en escalier à $(n+1)$ marches (voir pour le même exemple figure 2.5).

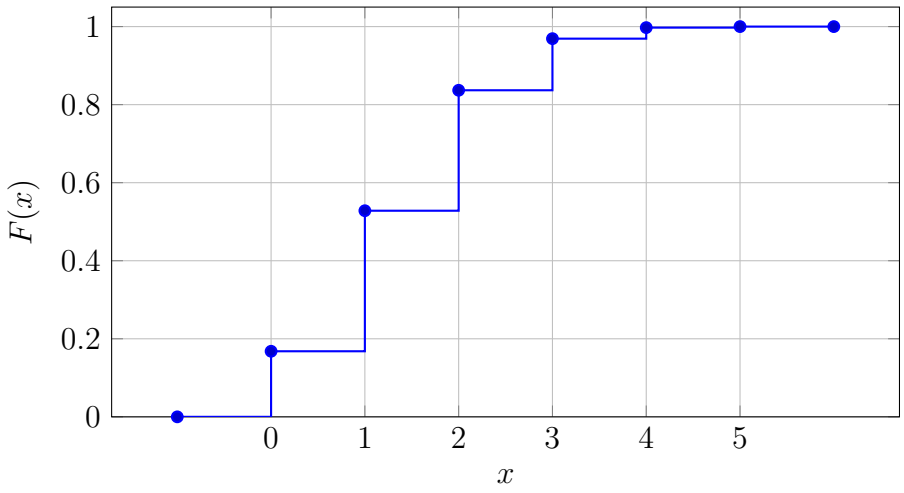


FIGURE 2.5 – Fonction de répartition de la loi binomiale $\mathcal{B}(5, 0.3)$.

Les moments $p = 1,\ 2$

L'espérance et la variance d'une variable $S_n \simeq \mathcal{B}(n,\ p)$ se calcule de deux manières comme suit

$$\mathbf{E}(S_n) = \sum_{k=0}^{n} k\mathbf{P}[S_n = k] = \sum_{k=0,n} kC_n^k p^k (1-p)^{(n-k)}$$

$$= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{(k-1)} (1-p)^{(n-k)}.$$

En utilisant le changement de variable suivant $k' = k - 1 \Leftrightarrow k = k' + 1$, on obtient

$$\mathbf{E}(S_n) = np \sum_{k'=0}^{n-1} \frac{(n-1)!}{k'!(n-k'-1)!} p^{k'} (1-p)^{(n-k'-1)}.$$

En réutilisant un changement de variable $n' = n - 1$, on obtient

$$\mathbf{E}(S_n) = np \underbrace{\sum_{k'=0}^{n'} \frac{n'!}{k'!(n'-k')!} p^{k'} (1-p)^{(n'-k')}}_{=1}$$

$$= np.$$

La deuxième méthode

$$\mathbf{E}(S_n) = \mathbf{E}\left(\sum_{k=1}^{n} X_k\right) \overset{\mathbf{E} \text{ linéaire}}{=} \sum_{k=1}^{n} \underbrace{\mathbf{E}(X_k)}_{=p} = np.$$

Et la variance est donnée par

$$\mathbf{Var}(S_n) = \mathbf{E}(S_n^2) - (\mathbf{E}(S_n))^2 = \sum_{i=0}^{n} k^2 \mathbf{P}[S_n = k] - (np)^2$$

$$= \sum_{k=0}^{n} k^2 \frac{n!}{k!(n-k)!} p^k (1-p)^{(n-k)} - (np)^2.$$

De la même manière, on obtient

$$\mathbf{Var}(S_n) = np \sum_{k'=0}^{n-1} (k'+1) \frac{(n-1)!}{k'!(n-k'-1)!} p^{k'} (1-p)^{(n-k'-1)} - (np)^2$$

$$= np \underbrace{\sum_{k'=0}^{n-1} \frac{(n-1)!}{k'!(n-k'-1)!} p^{k'} (1-p)^{(n-k'-1)}}_{=1}$$

$$+np \underbrace{\sum_{k'=0}^{n-1} k' \frac{(n-1)!}{k'!(n-k'-1)!} p^{k'} (1-p)^{(n-k'-1)} - (np)^2}_{=p(n-1)}$$

$$= np + n(n-1)p^2 - (np)^2 = np - np^2$$

$$= np(1-p).$$

La deuxième méthode pour la variance

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

$$\overset{\mathbf{Var} \text{ semi linéaire}}{=} \sum_{k=1}^{n} \mathbf{Var}(X_k)$$

$$= np(1-p).$$

En résumé $S_n \simeq \mathcal{B}(n,\ p) \Leftrightarrow$

1. $k \in \{0,\ 1, \ldots, n\}$,
2. $\mathbf{P}[S_n = k] = C_n^k p^k (1-p)^{(n-k)}$, avec $k = 0,\ 1, \ldots, n$,
3. $\mathbf{E}(S_n) = np$,
4. $\mathbf{Var}(S_n) = np(1-p)$.

La loi binomiale est utilisée que si les expériences sont **non exhaustives**, c'est la loi du tirage **avec remise** et les événements considérés doivent être **indépendants.**

2.6.3 Loi hypergéométrique

Cette loi de probabilité discrète est très proche de la loi binomiale. Elle décrit une suite d'épreuves dont le résultat aléatoire est binaire (de type succès, échec). La seule différence est qu'un individu, ne peut apparaître deux fois. Il s'agit donc d'un tirage **exhaustif**. En autre terme « **sans remise** ». En plus du paramètre p qui est la probabilité de l'événement favorable, la loi hypergéométrique fait intervenir deux paramètres de taille : celui de l'échantillon (n) et celui de la population de référence (N). C'est-à-dire, considérons une population de N objets parmi lesquels M sont d'un type (noté type I) et $N - M$ d'un second type (noté type II), dans laquelle en tire sans remise $n \leq N$ individus. Soit X une v.a. qui représente le nombre d'individus de type I (par exemple) parmi n et $p = \frac{M}{N}$ représente la probabilité de tirer un individu de type I. Remarquer que $\max(0,\ n - N + M) \leq k \leq \min(n,\ M)$) et on note cette loi par $X \simeq \mathcal{H}\left(N,\ n,\ p = \frac{M}{N}\right)$.

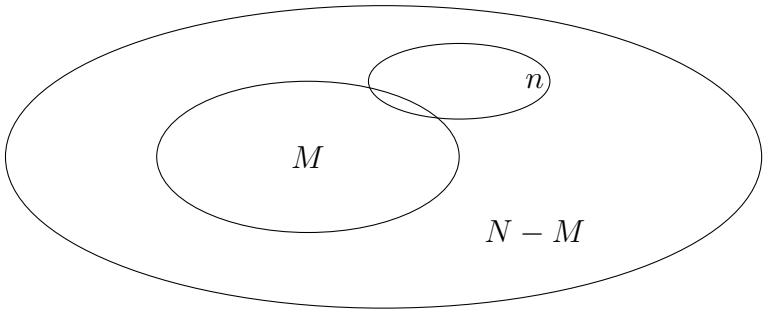


FIGURE 2.6 – Tirage sans remise de n individus

Loi de probabilité et moments

Pour une v.a. X de loi hypergéométrique alors on a,

$$\mathbf{P}[X = k] = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = p_k, \text{ avec } k \in \{\max(0,\ n-N+M),\ \ldots,\ \min(n,\ M)\}.$$

Cette loi est représentée par une fonction de répartition similaire à celle de la loi binomiale avec les sauts p_k. L'espérance est la même que celle de la loi binomiale, à savoir $\mathbf{E}(X) = np$. En revanche, la variance est un peu inférieure puisqu'à chaque tirage on retire une observation de l'échantillon, donc

$$\mathbf{Var}(X) = \frac{N-n}{N-1} np(1-p).$$

2.6.4 Loi géométrique

La loi géométrique est une loi de probabilité discrète qui représente la loi de l'observation du nombre d'épreuves de Bernoulli **identiques et indépendantes** qui se succèdent pour espérer **un premier succès.** Elle n'a donc qu'un paramètre, la probabilité de succès p. De cette probabilité découle celle d'un échec 1–p. Une variable aléatoire X suit une loi géométrique de paramètre p est notée par $X \simeq \mathcal{G}(p)$.

Loi de probabilité, fonction de répartition et moments

La loi géométrique est définie pour un n qui représente le nombre de tirages. La probabilité de remporter un premier succès à l'épreuve n est égale à

$$\mathbf{P}[X = n] = p(1-p)^{n-1} = p_n.$$

La fonction de répartition $F(n)$ définit la probabilité de réaliser au plus n épreuves pour obtenir le premier succès.

$$F_X(n) = \mathbf{P}[X \leq n] = 1 - (1-p)^n.$$

En effet, on a de $\mathbf{P}(X = n)$ que

$$\mathbf{P}[X \leq n] = \sum_{i=1}^{n} p(1-p)^i = p \sum_{i=0}^{n-1} (1-p)^i.$$

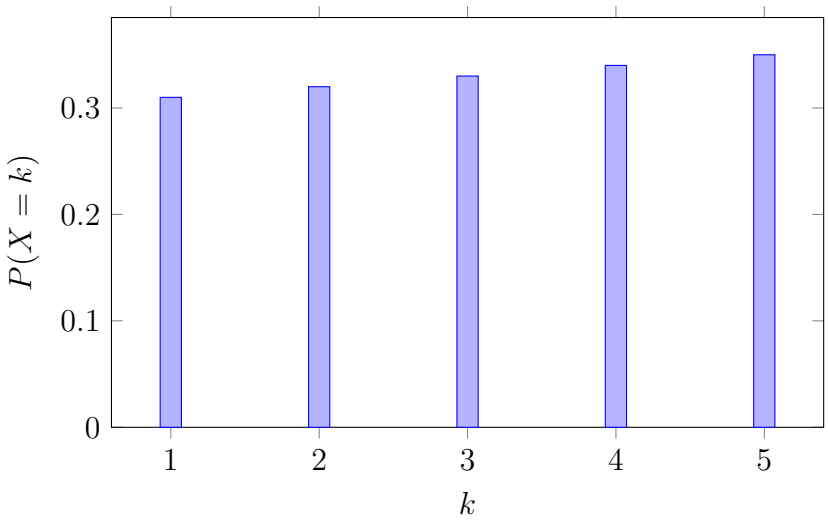


FIGURE 2.7 – Diagramme en bâtons de la loi géométrique ($p = 0.3$).

Qui représente la somme des n premiers termes d'une suite géométrique (d'où l'appellation) de raison $(1 - p)$ et de premier terme 1. Alors

$$\mathbf{P}[X \leq n] = p\frac{1 - (1 - p)^n}{1 - (1 - p)} = 1 - (1 - p)^n.$$

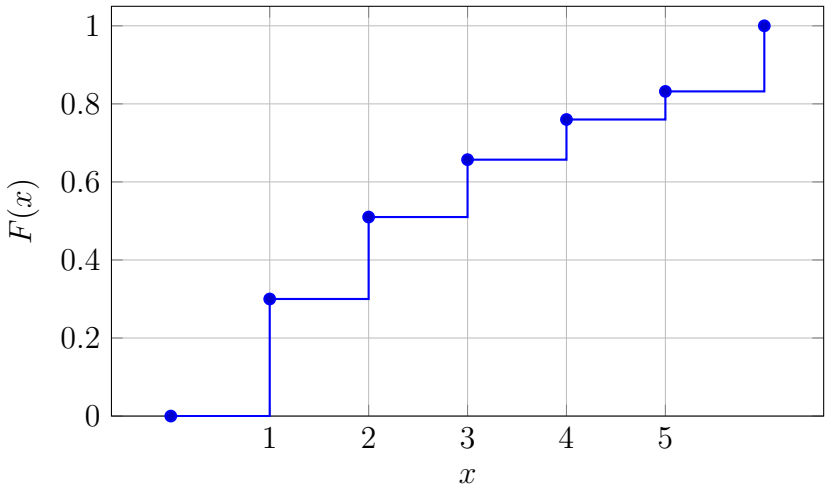


FIGURE 2.8 – Fonction de répartition de la loi géométrique ($p = 0.3$).

L'espérance est $\mathbf{E}(X) = \frac{1}{p}$ et la variance est donnée par $\mathbf{Var}(X) =$

$\frac{1-p}{p^2}$. Par exemple, il faut en moyenne six essais pour obtenir un 6 avec un dé cubique non truqué (l'inverse de 1/6). Encore une fois, attention à la problématique. Si par exemple on s'intéresse au nombre d'échecs en excluant le succès, nous avons : Temps d'attente. On lance une pièce de monnaie (truquée) dont la probabilité d'obtenir pile est p. On note X le nombre de lancers nécessaires pour obtenir pile. Alors X suit une loi géométrique de paramètre p.

2.6.5 Loi Poisson

La loi de poisson noté $\mathcal{P}(\lambda)$ est une loi de probabilité qui s'applique aux événements exceptionnels (une pandémie, une épidémie,...), elle décrit aussi le nombre de survenues d'un événement pendant une durée de temps déterminée. Elle survient de la même manière que la loi Binomiale mais avec une taille d'échantillon n assez grande et une probabilité de l'événement qui nous intéresse p assez petit. Le nombre aléatoire X de ces événements suit une loi de Poisson de paramètre $\lambda = np > 0$ qui représente la moyenne d'événements apparus.

Loi de probabilité, fonction de répartition et moments

Donc, pour un nombre réel positif $\lambda > 0$. On dit que la variable aléatoire X suit une loi de Poisson de paramètre λ, si X est une variable aléatoire discrète prenant la valeur entière n avec la probabilité

$$\mathbf{P}(X = n) = \frac{\exp^{-\lambda} \lambda^n}{n!} = p_n; \text{ si } n = 0; 1; 2; \ldots.$$

Voir par exemple la figure 2.9 pour $\lambda = 2$.

La fonction de répartition $F(n)$ est définie par

$$F(n) = \mathbf{P}[X \leq n] = \frac{\Gamma([n+1], \lambda)}{[n]!}, \ n \geq 0,$$

où $\Gamma(x, y)$ est la fonction gamma incomplète et où $[x]$ est la partie entière de x.

L'espérance de cette variable est $\mathbf{E}(X) = \lambda$ et sa variance est $\mathbf{Var}(X) = \lambda$.

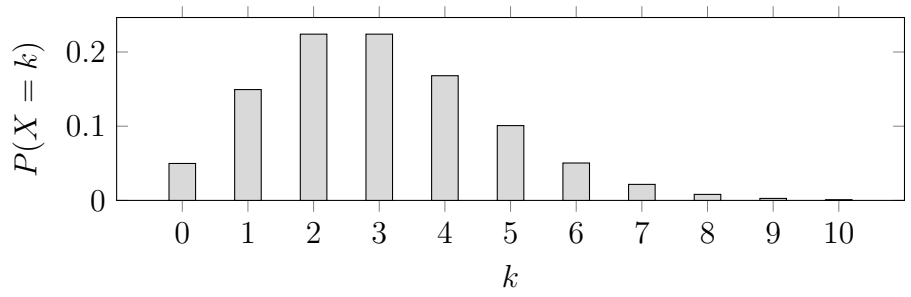


FIGURE 2.9 – Diagramme en bâtons de la loi de Poisson ($\lambda = 2$).

On a

$$\mathbf{E}(X) = \sum_{n\geq 0} n\mathbf{P}[X = n] = \sum_{n\geq 0} n\frac{\exp^{-\lambda}\lambda^n}{n!}.$$

On sait que

$$\sum_{n\geq 0} n\frac{\lambda^n}{n!} = \left[\lambda + 2\frac{\lambda^2}{2!} + 3\frac{\lambda^3}{3!} + \dots\right] = \lambda\left[1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots\right] = \lambda\sum_{n\geq 0}\frac{\lambda^n}{n!} = \lambda\exp^{\lambda}.$$

Donc $\mathbf{E}(X) = \lambda\exp^{\lambda}\exp^{-\lambda} = \lambda$. De même pour la variance, on a

$$\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X) = \sum_{n\geq 0} n\mathbf{P}[X = n] - \mathbf{E}^2(X)$$

$$= \sum_{n\geq 0} n^2\frac{\exp^{-\lambda}\lambda^n}{n!} - (\lambda^2).$$

Aussi

$$\sum_{n\geq 0} n^2\frac{\exp^{-\lambda}\lambda^n}{n!} = \exp^{-\lambda}\left[\lambda + 2\frac{\lambda^2}{1!} + 3\frac{\lambda^3}{2!} + \dots\right]$$

$$= \lambda\exp^{-\lambda}\left[1 + 2\frac{\lambda}{1!} + 3\frac{\lambda^2}{2!} + \dots\right]$$

$$= \lambda\exp^{-\lambda}\sum_{n\geq 0}(n+1)\frac{\lambda^n}{n!} = \lambda\exp^{-\lambda}\left[\underbrace{\sum_{n\geq 0} n\frac{\lambda^n}{n!}}_{=\lambda\exp^{\lambda}} + \sum_{n\geq 0}\underbrace{\frac{\lambda^n}{n!}}_{=\exp^{\lambda}}\right] = \lambda^2 + \lambda.$$

Donc $\mathbf{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

2.7 Approximation entre variables aléatoires discrètes

En pratique, dans quelques cas de figures, on se retrouve avec deux probabilités égales $\mathbf{P}(X = k) = \mathbf{P}(Y = k')$ de deux lois de probabilités différentes. D'où le terme approximation. Dans la suite, nous allons voir ses cas.

2.7.1 Approximation de la loi hypergéométrique par une loi Binomiale

Soit $X \simeq \mathcal{H}\left(N,\ n,\ p = \frac{M}{N}\right)$ avec N la taille de la population contenant M individus de type I et $N - M$ du type II. On tire de cette population un échantillon de taille n **sans remise**, avec $n \leq \inf(M, N - M)$. Lorsque $N \to \infty$, n et $p = \frac{M}{N}$ restent fixes (ce qui sous-entend que $M \to \infty$ tel que $\frac{M}{N} \to p$, on a

$$\lim_{\substack{N\to+\infty \\ p=\frac{M}{N}}} \mathbf{P}[X = k] = \mathbf{P}[Y = k].$$

Où $Y \simeq \mathcal{B}(n, p)$. Effectivement, d'après la définition de la loi de X

$$\mathbf{P}[X = k] = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n},\ \forall k = 0,\ \ldots, n$$

$$= \frac{\frac{M!}{k!(M-k)!}\ \frac{(N-M)!}{(n-k)!(N-M-n+k)!}}{\frac{N!}{n!(N-n)!}}$$

$$= \frac{n![M(M-1)\ldots(M-k+1)][(N-M)(N-M-1)\ldots(N-M-n+k+1)]}{k!(n-k)!N(N-1)\ldots(N-n+1)}$$

$$= C_n^k \frac{\overbrace{M(M-1)\ldots(M-k+1)}^{k\ termes}\ \overbrace{(N-M)(N-M-1)\ldots(N-M-n+k+1)}^{(n-k)\ termes}}{\underbrace{N(N-1)\ldots(N-n+1)}_{n\ termes}}.$$

On multiplie le numérateur et le dénominateur par $\frac{1}{N^n} = \frac{1}{N^k}\frac{1}{N^{n-k}}$, pour obtenir

$$\mathbf{P}[X = k] =$$

$$C_n^k \frac{\left[\frac{M}{N}\left(\frac{M}{N} - \frac{1}{N}\right)\ldots\left(\frac{M}{N} - \frac{(k-1)}{N}\right)\right]\left[\left(\frac{N}{N} - \frac{M}{N}\right)\left(\frac{(N-1)}{N} - \frac{M}{N}\right)\ldots\left(\frac{(N-n+k+1)}{N} - \frac{M}{N}\right)\right]}{\left[\frac{N}{N}\left(1 - \frac{1}{N}\right)\ldots\left(1 - \frac{(n-1)}{N}\right)\right]}.$$

Comme $p = \frac{M}{N}$ reste constant pour $N \to +\infty$, on a $\forall k = 0, \ldots, n$

$$\mathbf{P}[X = k] \to C_n^k p^k (1-p)^{(n-k)},$$

qui n'est autre que la loi Binomiale.

2.7.2 Approximation de la loi Binomiale par une loi de Poisson

Soit $X \simeq \mathcal{B}(n,\ p)$. Pour n assez grand et p assez petit, on approche la loi de X par une loi de Poisson de paramètre $\lambda = np$. On effet,
$\mathbf{P}(X = k) = C_n^k p^k (1-p)^{(n-k)}$

$$= \frac{1}{k!} n(n-1)\ldots(n-k+1) p^k (1-p)^{(n-k)}$$

$$= \frac{n^k}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{(k-1)}{n}\right) p^k (1-p)^{(n-k)}.$$

Aussi, on a $p = \frac{\lambda}{n}$ et $n \to \infty \Leftrightarrow p \to 0$, donc

$$\mathbf{P}(X = k) = \frac{\lambda^k}{k!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{(k-1)}{n}\right)\frac{\left(1 - \frac{\lambda}{n}\right)^n}{\left(1 - \frac{\lambda}{n}\right)^k}.$$

Comme

$$\left(1 - \frac{\lambda}{n}\right)^n \underset{n\to+\infty}{\longrightarrow} \exp^{-\lambda},$$

donc

$$\mathbf{P}(X = k) \underset{n\to+\infty}{\longrightarrow} \frac{\exp^{-\lambda} \lambda^k}{k!},$$

qui est la loi d'une v.a. r. de Poisson de paramètre $\lambda = np$.
En pratique, cette approximation s'applique dès que $n > 50$ et $p < 0.10$.

2.8 Exercices

Exercice 10 *Soient X une application de Ω dans $\mathbb{R}$, A, B deux sous-ensembles de $\mathbb{R}$ et $A_i \subset \mathbb{R}$, $\forall i \in I \subset \mathbb{N}$, alors*

1. $X^{-1}(\emptyset) = \emptyset$.
2. $X^{-1}(\mathbb{R}) = \Omega$.
3. $X^{-1}(C_{\mathbb{R}}B) = C_{\Omega}X^{-1}(B)$.
4. $X^{-1}(\cap_{i\in I}Ai) = \cap_{i\in I}X^{-1}(Ai)$.
5. $X^{-1}(\cup_{i\in I}Ai) = \cup_{i\in I}X^{-1}(Ai)$.

Exercice 11 *Soient $(\Omega, \mathcal{A}, \mathbf{P})$ e.p., X une v.a.r. Soit $\mathbf{P}_X$ définie sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ par*

$$\forall A \in \mathcal{B}(\mathbb{R});\ \mathbf{P}_X(A) = \mathbf{P}[\omega \in \Omega : X(\omega) \in A] = \mathbf{P}[X^{-1}(A)].$$

L'application $\mathbf{P}_X$ est une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Exercice 12 *Soient $(\Omega;\ \mathcal{A};\ P)$ un espace probabilisé et A un événement de $\mathcal{A}$. On appelle fonction indicatrice de l'événement A la fonction notée $\mathbf{1}_A$ (ou I_A) définie de $(\Omega;\ \mathcal{A})$ dans $(\mathbb{R}; \mathcal{B}(\mathbb{R}))$ par*

$$\mathbf{1}_A(\omega) = \begin{cases} 0 & si\,;\ \omega \notin A \\ 1 & si\,;\ \omega \in C_{\Omega}A \end{cases}.$$

Montrer que la fonction indicatrice $\mathbf{1}_A$ est une variable aléatoire.

Exercice 13 *Pour $\theta \in]0, 1[$, on définit la suite p_k par*

$$p_k = \begin{cases} C\theta \min\{k, 8-k\} & si\,;\ k = 1, \ldots, 7 \\ 1 - \theta & si\,;\ k = 8 \end{cases},$$

où C est une constante positive. Déterminer C de sorte que p_k soit une loi de probabilités sur $\{1, \ldots, 8\}$.

Exercice 14 *Soit X une variable aléatoire discrète à valeurs dans $\{1, 2, \ldots, n\}$ de loi $\mathbf{P}(X = i) = Ci \forall i \in \{1, 2, \ldots, n\}$ où C est une constante.*

1. *Déterminer la valeur de la constante C.*

2. *Déterminer la fonction de répartition de la loi de* X.
3. *Déterminer la loi de* $Y = n - X$.
4. *Déterminer la loi de* $Z = n + X$.

Exercice 15 *Soit* X *une variable aléatoire discrète*

1. *Soient* g *et* h *deux fonctions de* $\mathbb{R}$ *et* $a, b \in \mathbb{R}$, *montrer que*

$$\mathbf{E}(ag(X) + bh(X)) = a\mathbf{E}(g(X)) + b\mathbf{E}(h(X)).$$

2. *Aussi, pour* $a, b \in \mathbb{R}$, *montrer que*

$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b.$$

3. *Pour* $a \in \mathbb{R}$, *montrer que* $\mathbf{E}(a) = a$.
4. *Montrer que* $\mathbf{E}((X - \mathbf{E}(X))^1) = 0$, *et* $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}(X)^2$.
5. *Pour* $a \in \mathbb{R}$, *montrer que* $\mathbf{Var}(a) = 0$.
6. *Pour* $a \in \mathbb{R}$, *montrer que* $\mathbf{Var}(aX) = a^2\mathbf{Var}(X)$.

Exercice 16 *Soit* X *et* Y *deux variables aléatoires discrètes, indépendantes à valeurs dans* $\{0, 1\}$ *de lois respectives*

$$\mathbf{P}(X = 1) = p_1,\ \mathbf{P}(X = 0) = 1 - p_1,\ \mathbf{P}(Y = 1) = p_2,\ \mathbf{P}(Y = 0) = 1 - p_2.$$

On pose $U_1 = X + Y$ *et* $U_2 = XY$.

1. *Déterminer les lois de* U_1 *et de* U_2.
2. *Calculer* $\mathbf{E}[U_1],\ \mathbf{Var}[U_1],\ \mathbf{E}[U_2]$ *et* $\mathbf{Var}[U_2]$.

Exercice 17
Soient X *une v. a. r. et* $I \in \mathbb{R}$. *En utilisant les définitions montrer ce qui suit*

1. *La définition d'une v. a.* $\mathbf{P}(X \in I) = \mathbf{E}(\mathbf{1}_I(X))$.
2. *La fonction de répartition* $\mathbf{P}(X \leq x) = \mathbf{E}(\mathbf{1}_{]-\infty,\ x]}(X))$.
3. *La fonction poids* $\mathbf{P}(X = k) = \mathbf{E}(\mathbf{1}_X(k))$.

Exercice 18
Soit X *une v. a. à valeurs dans* $\mathbb{N}$. g_X *vérifie les propriétés suivantes*

1. *g_X est croissante sur $[0,\ 1]$ avec $g_X(0) = \mathbf{P}[X = 0]$ et $g_X(1) = \mathbf{P}[X = 1]$.*
2. *g_X est $\mathcal{C}^\infty(]-1,\ 1[)$ est ses dérivées "engendrent les probabilités" :*

$$g_X^{(n)} = n!\mathbf{P}[X = n],\ \forall n \in \mathbb{N}.$$

3. *Soit Y une v. a. discrète ; X et Y on même loi $\Leftrightarrow g_X = g_Y$.*
4. *Si $X(X-1)\dots(X-k+1)$ est intégrable $(k \in \mathbb{N}^*)$ alors*

$$\mathbf{E}\left(X(X-1)\dots(X-k+1)\right) = \lim_{s\nearrow 1} g_X^{(k)}(s).$$

Exercice 19 *Considérons l'expérience qui consiste à jeter trois fois une pièce, la probabilité de tomber sur "pile" étant p à chaque fois. Supposons qu'à chaque jet donnant "pile", on gagne un Dinar et qu'à chaque jet donnant "face" on en perde un. Intéressons nous au gain total, soit X cette quantité.*

1. *Donner l'espace des épreuves correspondant.*
2. *Donner la loi de X.*

Exercice 20 *Si $X \simeq \mathcal{G}(p)$ montrer que pour tous $t;\ s > 0$ on a*

$$\mathbf{P}(X > s + t | X > t) = P(X > s), \textit{propriété d'absence de mémoire.}$$

Exercice 21 *Une boite contient 8 composants parmi lesquels 2 sont défectueux. Trois composants sont pris au hasard et sans remise de la boite. Soit X le nombre de composants défectueux dans l'échantillon. Donner la loi de X, ainsi que $\mathbf{E}(X)$ et $\mathbf{Var}(X)$.*

Exercice 22 *Une maladie congénitale a une prévalence de 1% chez les nouveau-nés. Pour tout entier $n > 0$ on note S_n la variable aléatoire donnant le nombre de nouveau-nés malades sur un échantillon de taille n. Quelle loi suit cette variable aléatoire (justifier) ? Donner sa moyenne ainsi que sa variance. On enregistre dans une clinique 10 naissances durant un week-end. Quelle est la probabilité pour que deux au moins de ces nouveau-nés soient atteints de la maladie ?*

Exercice 23 *Dans un bureau de poste un guichet traite de deux catégories d'opérations : le retraits à vue et les dépôts d'objets. On fait les hypothèses suivantes : pendant tout intervalle de temps d'amplitude T exprimé en minutes et inclus dans les heurs d'ouverture, le nombre de personnes qui se présentent au guichet pour un retrait à vue est représenté par une v. a. r. X_a de loi de poisson de paramètre aT $(a > 0)$. Le nombre de personnes qui se présentent pour un dépôt est présenté par une v. a. r. X_b de loi de poisson de paramètre bT $(b > 0)$. Sachant que X_a et X_b sont indépendantes.*

1. *Montrer que la loi de la v. a. r. associée au nombre total de personnes qui se présentent au guichet est de loi de poisson.*
2. *Si a et b, quelle est la probabilité qu'aucun client ne se présente entre 10 heurs et 10h 5minutes.*
   Les employés terminent leur travail à 19h, le bureau de poste fermes ses portes à 18h 45min. A 18h 45min, il y a au guichet entre 18h35min et 18h 45min.
   Exprimer, en fonction de a, b, et n, la loi de probabilité associée au nombre de personnes de cette queue qui sont venues pour un retrait à vue.
3. *Pour $n = 5$, $a = 0.4$ et $b = 0.2$; calculer la probabilité pour que, dans cette queue, le nombre de personnes venues pour un retrait à vue soit inférieur au nombre de celles venues pour un dépôt.*

3 Variables aléatoires absolument continues

Dans ce chapitre, les v.a.r continues sont abordées d'une manière très explicite. La fonction de répartition ainsi que la densité avec les moments sont exposés. Deux nouvelles fonctions sont introduites, celle qui génère les moments et la caractéristique. Enfin, la présentation des lois de probabilité absolument continues usuelles achèvera cette partie.

3.1 Les v. a. r. absolument continus

Il faut bien comprendre la différence entre discret et continu. il s'agit du cardinal de l'ensemble fondamental Ω. S'il est fini ou dénombrable, après avoir fait un codage, on parle de v. a. r. discrète. Mais, lorsqu'il est infini (c-à-d, $\Omega \subseteq \mathbb{R}$), on obtient une v.a.r. continue. La représentation graphique de cette dernière est un histogramme donné par la figure (3.1) ou alors dans le cas dense la figure (3.2).

Remarquer qu'on peut pas calculer la somme des probabilités ponctuelles $\mathbf{P}[X = k]$ sur $[a,\ b]$, mais on peut considérer la surface comprise entre l'axe et la fonction f de la figure (3.3).

Donc, dans le cas continu, on ne parle plus de somme, de probabilité et de valeurs distinctes mais plutôt d'intégrale, de densité et d'intervalle de valeurs.

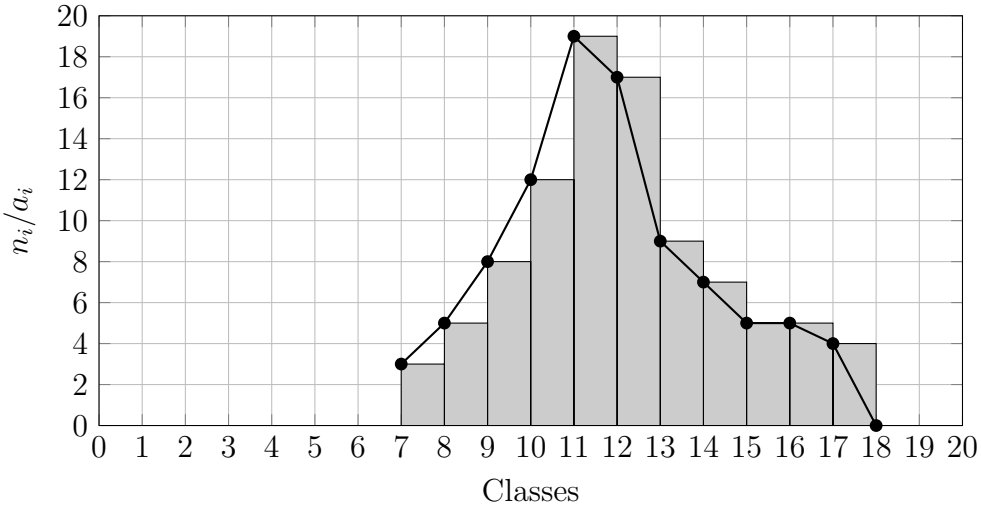


Figure 3.1 – Histogramme.

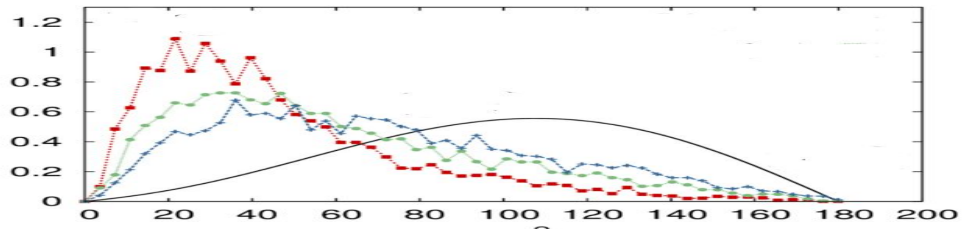


Figure 3.2 – Quatre fonctions de répartitions

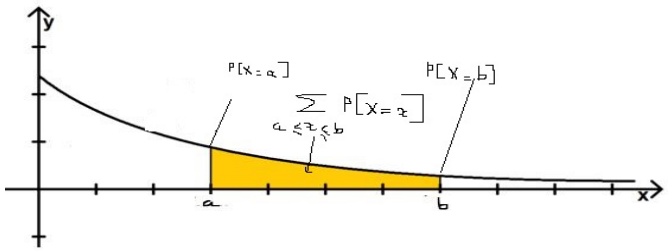


Figure 3.3 – Surface

3.1.1 Fonction de densité

D'une manière générale, si la probabilité qu'une variable aléatoire X appartienne à un intervalle peut s'écrire comme l'intégrale d'une fonction f sur cet intervalle, on dira que cette variable aléatoire admet la densité f.

Définition 5 *Densité*
Une fonction $f : \mathbb{R} \longrightarrow [0; 1[$, intégrable selon Riemann, s'appelle une densité si

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

La densité remplace la fonction masse, ce qui lui implique des propriétés spécifiques.

Propriété 2 *Soit f une fonction de densité d'une v.a.r continue X, alors*

1. $\mathbf{P}(a \leq X \leq b) = \mathbf{P}(a < X \leq b)$
   $= \mathbf{P}(a \leq X < b) = \mathbf{P}(a < X < b) = \int_a^b f(t)dt.$
2. $\mathbf{P}(t < X \leq t + dt) = f(t)dt.$
3. *La fonction définie par $\mathbf{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt$ calcule le cumule des probabilités $\sum_{k \leq x} \mathbf{P}(X = k)$ sur l'intervalle $]-\infty,\ x]$.*

3.2 Fonction de répartition

Une fonction $F : \mathbb{R} \longrightarrow [0; 1]$ est une fonction de répartition si

1. F est croissante : $x \leq y \Rightarrow F(x) \leq F(y)$.
2. F est continue à droite : $\lim\limits_{y \overset{>}{\to} x} F(y) = F(x),\ \forall x \in \mathbb{R}$.
3. $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$.

Une fonction de répartition F est dite absolument continue de densité f si

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

Cette dernière définition implique que F s'écrit comme suit

$$F(x) = \mathbf{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

Donc pour ne pas confondre, dans la suite, on note F_X la fonction de répartition et f_X la densité de la v.a.r continue X.

Propriété 3 *Soit X une v.a.r. absolument continue de densité f_X, on a*

1. $F'_X(x) = f_X(x) \Leftrightarrow \frac{dF_X(x)}{dx} = f_X(x)$.
2. $\int_a^b f_X(t)dt = F_X(b) - F_X(a)$.
3. $\mathbf{P}(X = a) = 0 \Leftrightarrow F_X$ *est continue en* a.

Preuve 1 *1. Directement.*

2. *Des propriétés de l'intégrale on a* $\int_a^b f_X(t)dt = \int_{-\infty}^b f_X(t)dt - \int_{-\infty}^a f_X(t)dt = F_X(b) - F_X(a)$.
3. $\overset{?}{\Rightarrow}$) *Supposons que* $\mathbf{P}(X = a) = 0$ *et montrons que* F_X *est continue en* a. *Comme* F_X *est continue à droite par définition, donc reste a démontrer la continuité à gauche. On a*

$$\lim_{x \overset{<}{\to} a} F_X(x) = \lim_{x \overset{<}{\to} a} \mathbf{P}(X \leq x) = \mathbf{P}(X \leq a-) = \mathbf{P}(X \leq a-) + \underbrace{\mathbf{P}(X = a)}_{=0} = \mathbf{P}(X \leq a).$$

$\overset{?}{\Leftarrow}$) *Supposons que* F_X *converge à gauche et on démontre que* $\mathbf{P}(X = a) = 0$. *D'après la continuité de* F_X *en* a,

$$\lim_{x \overset{<}{\to} a} F_X(x) = F_X(a) = \mathbf{P}(X \leq a),$$

et comme

$$\lim_{x \overset{<}{\to} a} F_X(x) = \mathbf{P}(X \leq a-).$$

Donc

$$\mathbf{P}(X \leq a-) = \mathbf{P}(X \leq a) \Rightarrow \mathbf{P}(X = a) = 0.$$

Comme dans le cas discret, les moments d'ordre un et deux serons introduits dans la suite.

3.3 Les moments

Soient X une variable aléatoire réelle définie sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ de densité f_X, on a le théorème suivant

Théorème 3 *Soit g une application réelle continue dans $\mathbb{R}$, alors si*

$$\int_{\mathbb{R}} |g(x)| f_X(x)dx < \infty \Rightarrow \mathbf{E}(g(X)) = \int_{\mathbb{R}} g(x) f_X(x)dx.$$

D'après ce théorème, pour $g_1(x) = x$ et $g_2(x) = x^2$ qui sont deux fonctions continues sur $\mathbb{R}$, on a

$$\mathbf{E}(X) = \int_{\mathbb{R}} x f_X(x)dx \text{ et } \mathbf{E}(X^2) = \int_{\mathbb{R}} x^2 f_X(x)dx.$$

De la linéarité de la fonction intégrale, on a la linéarité de la fonction espérance.

Propriété 4

1. *Soient g et h deux fonctions de $\mathbb{R}$ et $a, b \in \mathbb{R}$, on a*
$$\mathbf{E}(ag(X) + bh(X)) = a\mathbf{E}(g(X)) + b\mathbf{E}(h(X)).$$
2. *Aussi, pour $a, b \in \mathbb{R}$, on a*
$$\mathbf{E}(aX + b) = a\mathbf{E}(X) + b.$$
3. *Pour $a \in \mathbb{R}$, on a $\mathbf{E}(a) = a$.*
4. *Le moment centré d'ordre $k \in \mathbb{N}^*$ de la v. a. r. continue X est donné par*
$$\mathbf{E}((X - \mathbf{E}(X))^k).$$
Donc, on a $\mathbf{E}((X-\mathbf{E}(X))^1) = 0$ et $\mathbf{E}((X-\mathbf{E}(X))^2) = \mathbf{E}(X^2)-\mathbf{E}^2(X)$. D'où que
$$\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X).$$
5. *Pour $a \in \mathbb{R}$, on a :* $\mathbf{Var}(a) = \mathbf{E}(a^2) - \mathbf{E}^2(a) = a^2 - a^2 = 0$.
6. *Pour $a \in \mathbb{R}$, on a*
$$\mathbf{Var}(aX) = \mathbf{E}(a^2X^2) - \mathbf{E}^2(aX) = a^2\mathbf{E}(X^2) - a^2\mathbf{E}^2(X) = a^2\mathbf{Var}(X).$$
7. *L'écart-type de X est donné par $\sigma_X = \sqrt{\mathbf{Var}(X)}$.*

Définition 6 *Si $\mathbf{E}[X] = 0$ et $\mathbf{Var}[X] = 1$, on dit que X est centrée réduite. On remarque que lorsque $\mathbf{E}|X| < +\infty$ et $\mathbf{Var}[X] < +\infty$, alors la v.a.r. $\frac{X-\mathbf{E}[X]}{\sigma_X}$ est centrée réduite.*

3.3.1 Inégalités faisant intervenir les moments

Les moments permettent de donner une indication sur la dispersion d'une variable. Cette dernière peut par exemple être précisée à l'aide des inégalités suivantes.

Théorème 4 *(Inégalité de Markov)*
Si X est une v.a.r. positive, on a pour tout réel $a > 0$

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

Preuve 2 *On a*

$$\mathbf{E}[X] = \int_{\mathbb{R}^+} x d\mathbf{P}_X(x) = \int_{[0;a[} x d\mathbf{P}_X(x) + \int_{[a;+\infty[} x d\mathbf{P}_X(x).$$

D'où

$$\mathbf{E}[X] \geq \int_{[a;+\infty[} x d\mathbf{P}_X(x) \geq a \int_{[a;+\infty[} d\mathbf{P}_X(x) = \mathbf{P}(X \geq a).$$

Théorème 5 *(Inégalité de Bienaymé-Chebychev)*
Si $\mathbf{E}[X^2] < +\infty$, alors on a pour tout réel $a > 0$

$$\mathbf{P}(|X - \mathbf{E}[X]| > a) \leq \frac{\mathbf{Var}[X]}{a^2}.$$

Preuve 3 *Il suffit d'appliquer l'inégalité de Markov à la variable aléatoire $(X - \mathbf{E}[X])^2$.*

Comme pour le cas d'une variable aléatoire discrète, la fonction de répartition représente une des méthodes qui permettent de caractériser la loi d'une variable aléatoire, mais il existe d'autres méthodes comme par exemple la fonction caractéristique. Cette dernière est une fonction plus pratique d'utilisation et nous allons la voir dans la section suivante.

3.4 Fonction caractéristique

Soit X une v. a. On appelle fonction caractéristique de X la fonction $\varphi_X : \mathbb{R} \longrightarrow \mathbb{C}$, définie par

$$\varphi_X(t) = \mathbf{E}(e^{itX}), \ \forall t \in \mathbb{R}.$$

Suivant la loi de X la formule s'écrit

— Si X est une v. a. discrète finie alors,

$$\varphi_X(t) = \sum_{j=1}^{n} e^{itx_j} \mathbf{P}[X = x_j],\ \forall t \in \mathbb{R}.$$

Comme la somme est finie alors, φ_X est bien définie.

— Si X est une v. a. discrète dénombrable alors,

$$\varphi_X(t) = \sum_{j=1}^{\infty} e^{itx_j} \mathbf{P}[X = x_j],\ \forall t \in \mathbb{R}.$$

On a $\forall t \in \mathbb{R}$,

$$|\varphi_X(t)| \leq \sum_{j=1}^{\infty} |e^{itx_j}| \mathbf{P}[X = x_j] = \sum_{j=1}^{\infty} \mathbf{P}[X = x_j] = 1. \tag{3.1}$$

— Si X est une v. a. continue de densité f_X alors,

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f_X(x) dx,\ \forall t \in \mathbb{R}.$$

Aussi, on a $\forall t \in \mathbb{R}$

$$|\varphi_X(t)| \leq \int_{-\infty}^{+\infty} |e^{itx}| f_X(x) dx = \int_{-\infty}^{+\infty} f_X(x) dx = 1. \tag{3.2}$$

Proposition 3 *Soit X une v. a., a et b deux réels. Alors la fonction caractéristique vérifie ce qui suit*

1. $\forall t \in \mathbb{R},\ |\varphi_X(t)| \leq 1.$
2. $\varphi_X(0) = 1.$
3. $\forall t \in \mathbb{R},\ \varphi_X(-t) = \overline{\varphi_X(t)}.$
4. $\forall t \in \mathbb{R},\ \varphi_{aX+b}(t) = e^{itb} \varphi_X(at).$
5. φ_X *est continue sur* $\mathbb{R}$.
6. *Soient* $X_1, \ldots, X_n$ *n v. a. indépendantes. Soit* $S_n = \sum_{j=1}^{n} X_j$. *Alors*

$$\varphi_{S_n}(t) = \prod_{j=1}^{n} \varphi_{X_j}(t),\ \forall t \in \mathbb{R}.$$

Preuve 4 *1. Cette propriété est bien démontrer par les relations (3.1) et (3.2).*

2. $\varphi_X(0) = \mathbf{E}(e^{i0X}) = \mathbf{E}(1) = 1$.
3. $\varphi_X(-t) = \mathbf{E}(e^{-itX}) = \overline{\varphi_X(t)}$.
4. $\varphi_{aX+b}(t) = \mathbf{E}(e^{it(aX+b)}) = \mathbf{E}(e^{itb}e^{it(aX)}) = e^{itb}\varphi_X(at)$.
5. *La continuité de cette fonction est un résultat directe de la continuité des fonctions somme et intégrale.*
6. $\varphi_{S_n}(t) = \mathbf{E}\left(e^{itS_n}\right) = \mathbf{E}\left(e^{it\sum_{j=1}^n X_j}\right) = \mathbf{E}\left(\prod_{j=1}^n e^{itX_j}\right)$

$$\overset{indépendance}{=} \prod_{j=1}^n \mathbf{E}\left(e^{itX_j}\right) = \prod_{j=1}^n \varphi_{X_j}(t).$$

Le théorème suivant permet d'avoir les plus importants résultats concernant la fonction caractéristique et le fait qu'elle définie la loi de probabilité d'une v. a..

Théorème 6 — *Si la fonction caractéristique de X est bien définie alors, la loi de probabilité est bien connue.*

— *Si X et Y deux v. a. tel que $\varphi_X = \varphi_Y \Rightarrow \mathbf{P}_X = \mathbf{P}_Y$.*

— *Si X une v. a. de fct caractéristique φ_X intégrable $\left(\int_{\mathbb{R}} |\varphi_X(t)|dt < \infty\right)$ alors, X admet une fct de densité f_X définie sur $\mathbb{R}$ par*

$$f_X(x) = \int_{\mathbb{R}} e^{-itx}\varphi_X(t)dt, \ \forall x \in \mathbb{R}.$$

La fonction caractéristique permet aussi d'engendré les moments d'ordre k (c-à-d $\mathbf{E}(X^k)$) d'une v.a.

Proposition 4 *Soit X une v. a. de densité f_X avec $\mathbf{E}(X^k) < \infty$, pour $k \in \mathbb{N}^*$ alors*

$$\mathbf{E}(X^k) = \frac{1}{i^k}\varphi_X^{(k)}(0).$$

Preuve 5 *On suppose que X admet une densité f_X. Si $E(X^k)$ existe, la fonction $x \longrightarrow (ix)^k e^{itx} f(x)$ est uniformément intégrable et d'après les propriétés des intégrales, φ_X est k fois dérivable et*

$$\varphi_X^{(k)}(t) = \int_{-\infty}^{+\infty} (ix)^k e^{itx} f_X(x)dx,$$

c'est-à-dire, pour tout $t \in \mathbb{R}$

$$\varphi_X^{(k)}(t) = \mathbf{E}\left((iX)^k e^{itX}\right).$$

En particulier, pour $t = 0$,

$$\varphi_X^{(k)}(0) = \mathbf{E}(i^k X^k e^0) = i^k \mathbf{E}(X^k).$$

Dans certains cas où les calculs directs sont complexes, cette proposition permet d'obtenir rapidement $\mathbf{E}(X)$ *et* $\mathbf{E}(X^2)$ *par*

$$\mathbf{E}(X) = -i\varphi'_X(0) \text{ et } \mathbf{E}(X^2) = -\varphi''_X(0).$$

Dans ce qui suit, on présente les lois de probabilités continues les plus utilisées dans la pratique. Ainsi que leur fonctions de répartition et les deux moments.

3.5 Lois de probabilités absolument continues usuelles

3.5.1 Loi uniforme

Une v.a X suit la loi uniforme sur l'intervalle $[a,\ b]$ si elle admet pour densité la fonction (voir figure(3.4))

$$f_X(x) = \begin{cases} \frac{1}{(b-a)} & \text{si}\,;\ a \leq x \leq b \\ 0 & \text{sinon} \end{cases}.$$

Et on écrit $X \simeq \mathcal{U}([a,\ b])$. L'espérance et la variance sont

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

Sa fonction de répartition est

$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } \ a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}.$$

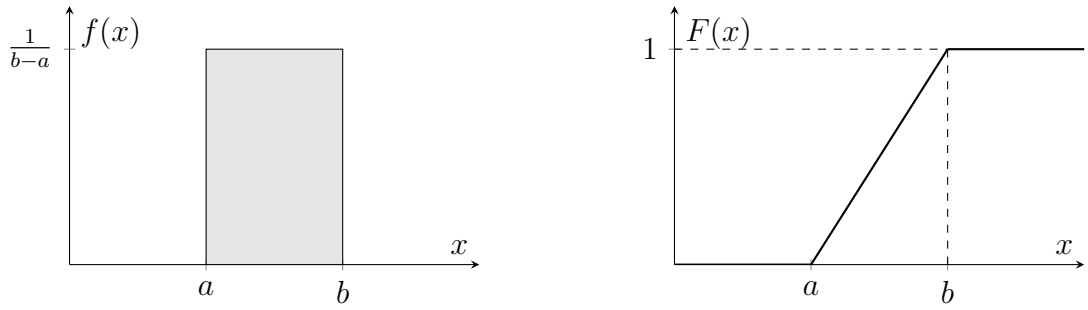


Figure 3.4 – La loi uniforme

3.5.2 Loi exponentielle

Une v.a X suit la loi exponentielle de paramètre $\lambda > 0$ si elle admet pour densité la fonction (voir figure(3.5))

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

et on écrit $X \simeq \mathcal{E}(\lambda)$. L'espérance et la variance de X sont donnés par

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

Sa fonction de répartition (voir figure(3.5)) est

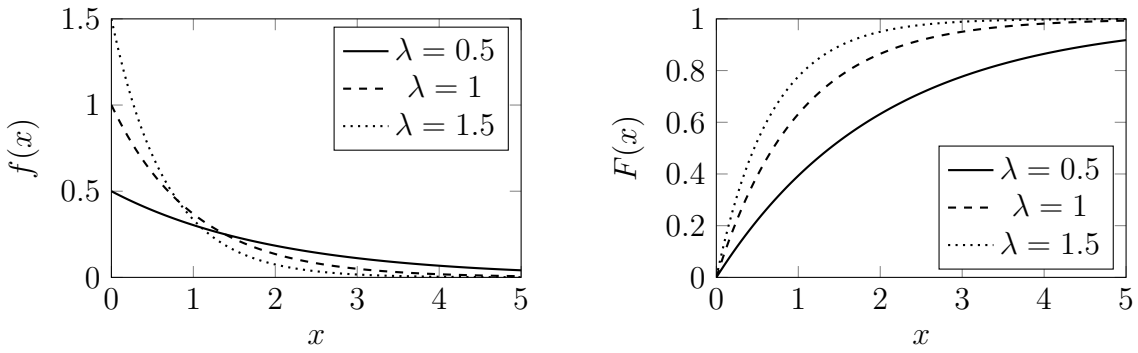


Figure 3.5 – Loi exponentielle

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

La loi exponentielle sert de modèle dans les problèmes de files d'attentes et de durée de vie.

3.5.3 Loi Gamma

Une v.a X suit la loi Gamma de paramètre $(a;\lambda)$ si elle admet pour densité la fonction (voir figure(3.6))

$$f_X(x) = \begin{cases} \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{[-\lambda x]} & \text{si } x > 0,\ \lambda \in \mathbb{R}_+^* \text{ et } a \in \mathbb{R}_+^* \\ 0 & \text{si } x \leq 0 \end{cases},$$

avec $\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{[-x]} dx$ et on écrit $X \simeq \Gamma(a;\ \lambda)$.

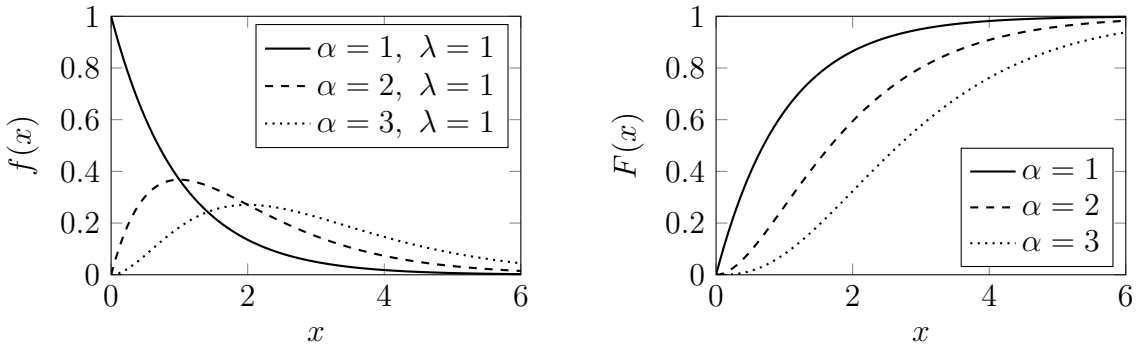


FIGURE 3.6 – Loi Gamma

L'espérance et la variance de X sont

$$\mathbf{E}(X) = \frac{a}{\lambda} \text{ et } \mathbf{Var}(X) = \frac{a}{\lambda^2}.$$

La fonction de répartition n'est pas facile à calculer mes par simulation sa représentation graphique est donnée par la figure (3.6).
Lorsque $a = 1$, on constate que la loi $\Gamma(a;\ \lambda)$ est la loi exponentielle de paramètre λ.

3.5.4 Loi Beta

Soient α et β deux réels positifs. On appelle loi Beta de paramètres $(\alpha,\ \beta)$ la loi de probabilité absolument continue dont la densité est définie

par

$$f_X(x) = \begin{cases} \frac{x^{\alpha-1}}{B(\alpha,\ \beta)}(1-x)^{\beta-1} & \text{si } x \in]0,\ 1[\\ 0 & \text{sinon} \end{cases}.$$

Cette mesure est identifiée par la notation Beta $(\alpha,\ \beta)$. (Voir figure (3.7)). Le facteur de normalisation $B(\alpha,\ \beta)$ apparaissant dans la formule précédente est la valeur en a et b de la fonction B d'Euler est donnée par

$$B(\alpha,\ \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Sa fonction de répartition de cette loi est donnée par

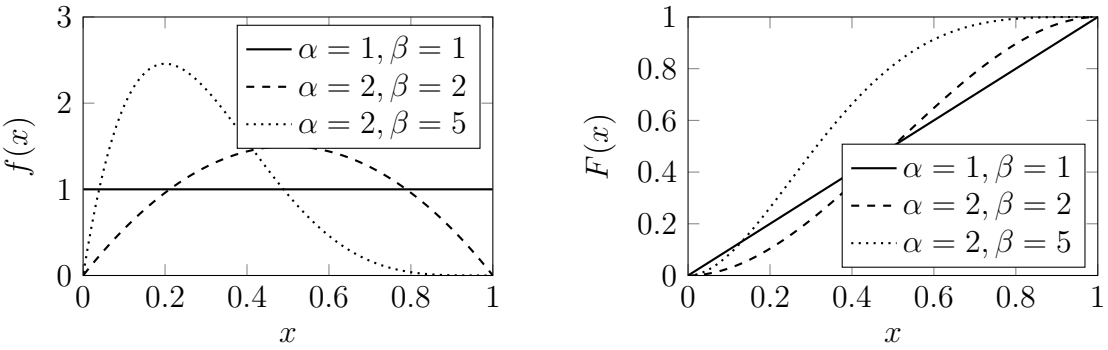


Figure 3.7 – Loi Beta

$$F_X(x) = \frac{Beta(\alpha,\ \beta)}{B(\alpha,\ \beta)}.$$

3.5.5 La loi normale

En pratique, la loi normale représente la loi la plus importante des lois de probabilité continues. Elle représente une modélisation d'un phénomène naturel comme par exemple le corps humain. La naissance, le développement ou la croissance, la stabilité et enfin la décroissance qui représente la vieillesse.

Loi normale (de Gauss, de Laplace)

On dit que la variable aléatoire X suit une loi normale de moyenne μ et de variance σ^2 si sa densité a pour expression

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\},\ \forall x \in \mathbb{R},\ \text{et } \sigma > 0,$$

et on note $X \simeq \mathcal{N}(\mu, \sigma^2)$.

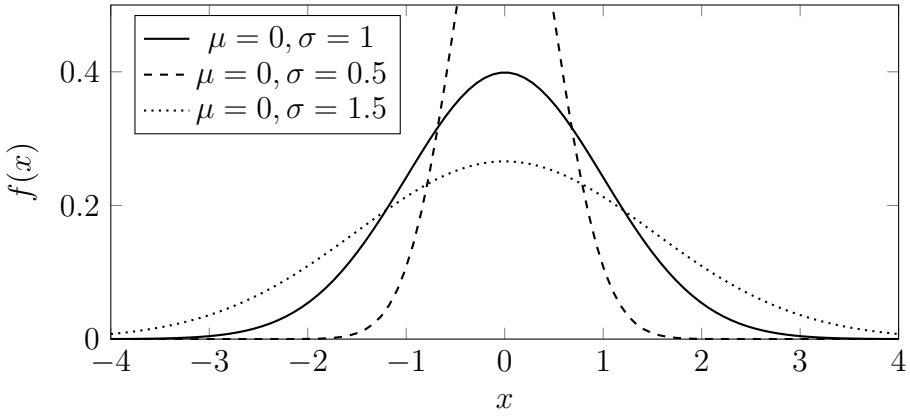


FIGURE 3.8 – Densité de la loi Normale

On vérifie que f_X est une densité de probabilité

1. $f_X(x) > 0 \ \ \forall x \in \mathbb{R}$ (évident).
2. $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$?
   On pose $t = \frac{x-\mu}{\sigma}$, $\ dt = \frac{dx}{\sigma} \Longrightarrow dx = \sigma dt$, alors

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} \sigma dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt = 1.$$

L'espérance mathématique et la variance de la v.a X sont données par

— L'espérance

$$\mathbf{E}(X) = \int_{-\infty}^{+\infty} x\ f_X(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

On pose $t = \frac{x-\mu}{\sigma}$, $x = \sigma t + \mu$, $dx = \sigma dt$ donc

$$\begin{aligned}
\mathbf{E}(X) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma t + \mu)\, e^{-\frac{1}{2}t^2}\ \sigma dt \\
&= \frac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{+\infty} \sigma t e^{-\frac{1}{2}t^2}\ dt + \mu \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt \right\} \\
&= \frac{1}{\sqrt{2\pi}} \left[-\sigma\ e^{-\frac{1}{2}t^2} \right]_{-\infty}^{+\infty} + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt = \mu,
\end{aligned}$$

car $\left[-\sigma e^{-\frac{1}{2}t^2} \right]_{-\infty}^{+\infty} = 0$ et $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt = 1$, d'où

$$\mathbf{E}(X) = \mu.$$

— La variance $\mathbf{Var}(X) = \mathbf{E}(X^2) - \mathbf{E}^2(X)$.

$$\begin{aligned}
\mathbf{E}\left(X^2\right) &= \int_{-\infty}^{+\infty} x^2\ f_X(x)\, dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2\ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&\overset{t=\frac{x-\mu}{\sigma}}{=} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(\sigma^2 t^2 + 2\mu\sigma t + \mu^2\right) e^{-\frac{1}{2}t^2}\ \sigma dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sigma^2 t^2 e^{-\frac{1}{2}t^2}\ dt + \frac{2\mu\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-\frac{1}{2}t^2}\ dt \\
&+ \frac{\mu^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2}\ dt \\
&= \underbrace{\frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t.t\ e^{-\frac{1}{2}t^2}\ dt}_{\text{intégration par partie}} + \frac{2\mu\sigma}{\sqrt{2\pi}} \left[-e^{-\frac{1}{2}t^2} \right]_{-\infty}^{+\infty} + \mu^2 \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left[-t e^{-\frac{1}{2}t^2} \right]_{-\infty}^{+\infty} + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2}\ dt + \mu^2
\end{aligned}$$

$= \sigma^2 + \mu^2$, donc $\mathbf{Var}(X) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$.

Remarque 7 *On peut calculer la variance de la v.a X en utilisant la*

définition de la variance comme suit

$$\begin{aligned}
\mathbf{Var}(X) &= \mathbf{E}\left[(X-\mathbf{E}(X))^2\right] = \mathbf{E}\left[(X-\mu)^2\right] \\
&= \int_{-\infty}^{+\infty} (x-\mu)^2 \ f(x) \ dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} (x-\mu)^2 \ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&= \frac{\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(\frac{x-\mu}{\sigma}\right)^2 \ e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\
&\overset{y=\frac{x-\mu}{\sigma}}{=} \underbrace{\frac{\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 \ e^{-\frac{1}{2}y^2} \sigma dy}_{\text{intégration par partie}}.
\end{aligned}$$

Donc $\mathbf{Var}(X) = \sigma^2$.

La fonction de répartition F_X est définie comme suit

$$\begin{aligned}
F \ : \ & \mathbb{R} \longrightarrow [0,1] \\
& x \longrightarrow F_X(x) = \int_{-\infty}^{x} f_X(t) \ dt.
\end{aligned}$$

Cette fonction est très difficile à calculer, donc il est judicieux de trouver une autre méthode pour la calculer. Aussi avant dont parler on introduit la partie ci dessous.

Loi normale centrée réduite

On dit que la variable aléatoire Z suit une loi normale centrée réduite (standard) si sa densité de probabilité a pour expression

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}, \ \forall z \in \mathbb{R}.$$

Et on note $Z \simeq \mathcal{N}(0,1)$.

D'après la définition de la loi normale, on remarque que la densité $f_Z = f_X$ pour $\mu = 0$ et $\sigma = 1$. Donc f_Z est bien une densité de probabilité et en

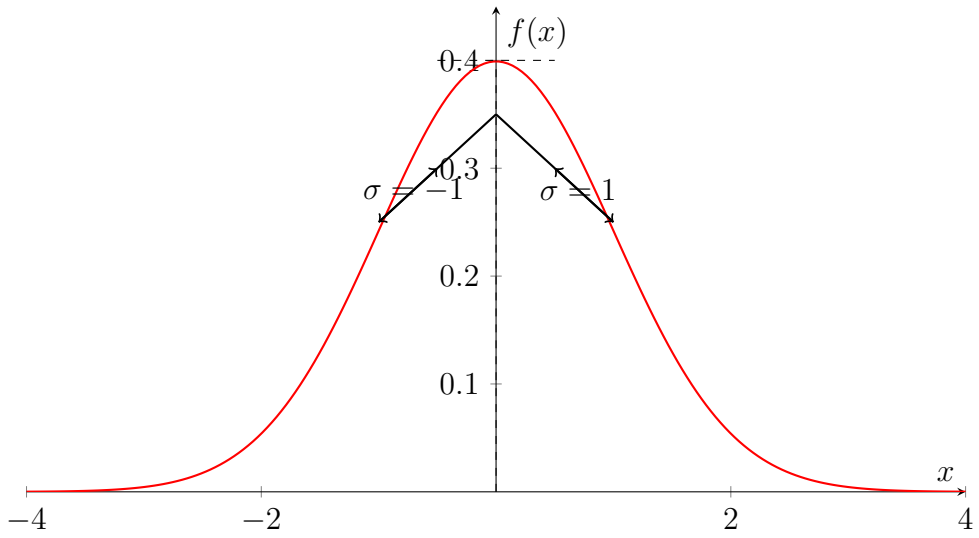


Figure 3.9 – Densité de la loi Normale centré réduite

calculant son espérance mathématique et sa variance on obtient

$$\begin{aligned}\mathbf{E}(Z) &= \int_{-\infty}^{+\infty} z f(z)\, dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z\ e^{-\frac{1}{2}z^2} dz \\ &= \frac{1}{\sqrt{2\pi}} \left[-e^{-\frac{1}{2}z^2}\right]_{-\infty}^{+\infty} = 0.\end{aligned}$$

donc $\mathbf{E}(Z) = 0$, d'où le nom centrée.

$$\begin{aligned}\mathbf{Var}(Z) &= \mathbf{E}\left(Z^2\right) + \underbrace{\mathbf{E}^2(Z)}_{=0} \\ &= \mathbf{E}\left(Z^2\right) \\ &= \int_{-\infty}^{+\infty} z^2 f_Z(z)\, dz \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^2\ e^{-\frac{1}{2}z^2} dz.\end{aligned}$$

Par intégration par parties $U = z \Rightarrow dU = dz$ et $dV = ze^{-\frac{1}{2}z^2}dz \Rightarrow V =$

$-e^{-\frac{1}{2}z^2}$. On obtient

$$\mathbf{E}\left(Z^2\right) = \frac{1}{\sqrt{2\pi}}\left\{\left[-z\ e^{-\frac{1}{2}z^2}\right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2}dz\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2}dz = 1,$$

puisque $\left[-z\ e^{-\frac{1}{2}z^2}\right]_{-\infty}^{+\infty} = 0$ et $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}z^2}dz = 1$. Donc $\mathbf{Var}\left(Z\right) = 1$.
La fonction de répartition de la v.a Z définie par

$$F_Z(t) = \mathbf{P}\left(Z \leq t\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{1}{2}z^2}dz,$$

admet des propriétés assez pratiques cités ci dessous.

Propriété 5 *1.* $F_Z\left(0\right) = \frac{1}{2}$.

2. $\forall t \in \mathbb{R},\ F_Z\left(t\right) + F_Z\left(-t\right) = 1 \Longrightarrow F_Z\left(-t\right) = 1 - F_Z\left(t\right).$

3. $\forall t \in \mathbb{R},\ F_Z\left(t\right) - F_Z\left(-t\right) = 2F_Z\left(t\right) - 1.$

4. $\mathbf{P}\left(Z \leq -t\right) = \mathbf{P}\left(Z > t\right)$ *de la symétrie de* f_Z.

5. $\forall\ a, b \in \mathbb{R};\ (a < b)$ *on a* $\mathbf{P}\left(a \leq Z \leq b\right) = F_Z\left(b\right) - F_Z\left(a\right).$

Remarque 8 *Si la v.a* $X \simeq \mathcal{N}\left(\mu, \sigma^2\right)$ *alors la v.a* $Z = \frac{X-\mu}{\sigma}$ *suit la loi normale centrée réduite* $\left(Z \simeq \mathcal{N}\left(0,1\right)\right)$.

Pour travailler avec cette loi et à l'aide de programme informatique, il a été possible de calculer la fonction de répartition de la loi normale centrée réduite. Les résultats on été regroupés dans un tableau (voir appendice 5.3). Dans la suite on traite des exemples pour voir comment utiliser ce tableau.

Exemple 1 *Soit* $X \simeq \mathcal{N}\left(\mu, \sigma^2\right)$, *calculer* $F_X(a)$ *où* $a \in \mathbb{R}$.

$$\begin{aligned} F_X\left(a\right) &= \mathbf{P}\left(X \leq a\right) \\ &= \mathbf{P}\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) \\ &= \mathbf{P}\left(Z \leq \frac{a-\mu}{\sigma}\right) = F_Z\left(\frac{a-\mu}{\sigma}\right) \ \text{où}\ Z \simeq \mathcal{N}\left(0,1\right). \end{aligned}$$

Exemple 2 *Soit* $X \simeq \mathcal{N}\left(3,9\right)$. *Calculer*

1. $\mathbf{P}\left(2 < X < 5\right).$

2. $\mathbf{P}(X > 0)$.
3. $\mathbf{P}(|X - 3| > 6)$.

On a

1. $\mathbf{P}(2 < X < 5) = \mathbf{P}\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right) \overset{Z \simeq \mathcal{N}(0,1)}{=} \mathbf{P}\left(\frac{-1}{3} < Z < \frac{2}{3}\right)$

$$= F_Z\left(\frac{2}{3}\right) - F_Z\left(-\frac{1}{3}\right) = F_Z\left(\frac{2}{3}\right) - \left[1 - F_Z\left(\frac{1}{3}\right)\right] = 0.3779.$$

2. $\mathbf{P}(X > 0) = \mathbf{P}\left(\frac{X-3}{3} > \frac{0-3}{3}\right) \overset{Z \simeq \mathcal{N}(0,1)}{=} \mathbf{P}(Z > -1)$

$$= 1 - \mathbf{P}(Z \leq 1) = 1 - F_Z(-1) = 1 - [1 - F_Z(1)] = F_Z(1) = 0.8413.$$

3. $\mathbf{P}(|X - 3| > 6) = 1 - P(|X - 3| \leq 6) = 1 - P(-6 \leq X - 3 \leq 6)$

$$= 1 - P(-3 \leq X \leq 9) \overset{Z \simeq \mathcal{N}(0,1)}{=} 1 - P(-2 \leq Z \leq 2)$$

$$= 1 - [F_Z(2) - F_Z(-2)] = 1 - [F_Z(2) - (1 - F_Z(2))] = 2 - 2F_Z(2) = 0.0456.$$

Exemple 3 *Calculer la surface comprise entre le graphe de la loi standard et les droites*

1. $z = 0,\ z = 0.5$.
2. $z = -2.24,\ z = 1.12$.
3. $z = 1,\ z = 2$.

On a

1. $\mathbf{P}(0 \leq Z \leq 0.5) = F_Z(0.5) - F_Z(0) = 0.1915$, *d'après la table de la loi* $\mathcal{N}(0, 1)$.
2. $\mathbf{P}(-2024 \leq Z \leq 1.12) = F_Z(1.12) - F_Z(-2.24)$

$$= \quad F_Z(1.12) - [1 - F_Z(2.24)] = 0.8686 - 1 + 0.9875 = 0.8561.$$

3. $\mathbf{P}(1 \leq Z \leq 2) = F_Z(2) - F_Z(1) = 0.9772 - 0.8413 = 0.1359$.

La proposition suivante permet la maîtrise linéaire des caractéristiques de la loi normale.

Proposition 5 *Soit* $(a, b) \in \mathbb{R}^* \times \mathbb{R}$.

1. *Si* $Z \simeq \mathcal{N}(0,1) \Longrightarrow Y = aZ + b \simeq \mathcal{N}(b, a^2)$.
2. *Si* $X \simeq \mathcal{N}(\mu, \sigma^2) \Longrightarrow Y = aX + b \simeq \mathcal{N}(a\mu + b, a^2\sigma^2)$.

Preuve 6 *1. Soit* F_Y *la fonction de répartition de la v.a* $Y = aZ + b$.

$$F_Y(y) = P(Y \leq y) = P(aZ + b \leq y) = P(aZ \leq y - b)$$

— *Cas où* $a > 0$

$$F_Y(y) = P\left(Z \leq \frac{y-b}{a}\right) = F_Z\left(\frac{y-b}{a}\right),$$

où F_Z *est la fct de répartition de la v.a* Z, *donc*

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_Z\left(\frac{y-b}{a}\right) = \frac{1}{a}f_Z\left(\frac{y-b}{a}\right) \\ &= \frac{1}{a}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-b}{a}\right)^2}. \end{aligned}$$

— *Cas où* $a < 0$

$$\begin{aligned} F_Y(y) &= P(aZ \leq y - b) = P\left(Z \geq \frac{y-b}{a}\right) \\ &= 1 - P\left(Z < \frac{y-b}{a}\right) = 1 - F_Z\left(\frac{y-b}{a}\right) \end{aligned}$$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left[1 - F_Z\left(\frac{y-b}{a}\right)\right] = -\frac{1}{a}f_Z\left(\frac{y-b}{a}\right) \\ &= -\frac{1}{a}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-b}{a}\right)^2}. \end{aligned}$$

Donc pour $a \neq 0$ *on obtient*

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-b}{a}\right)^2},\ \forall y \in \mathbb{R} \\ &\Longrightarrow\ Y \simeq \mathcal{N}\left(b, a^2\right). \end{aligned}$$

2. Soit F_Y *la fonction de répartition de la v.a* $Y = aX + b$

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(aX + b \leq y) = \mathbf{P}(aX \leq y - b).$$

— *Cas où* $a > 0$

$$F_Y(y) = \mathbf{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

Donc

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y-b}{a}\right) = \frac{1}{a}f_X\left(\frac{y-b}{a}\right) \\ &= \frac{1}{a\ \sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\frac{y-b}{a}-m}{\sigma}\right)^2} = \frac{1}{a\ \sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-b-am}{a\sigma}\right)^2}. \end{aligned}$$

— *Cas où* $a < 0$

$$F_Y(y) = P\left(X \geq \frac{y-b}{a}\right) = 1 - P\left(X < \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right)$$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left[1 - F_X\left(\frac{y-b}{a}\right)\right] = -\frac{1}{a}f_X\left(\frac{y-b}{a}\right) \\ &= -\frac{1}{a\ \sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\frac{y-b}{a}-m}{\sigma}\right)^2} = -\frac{1}{a\ \sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-b-a\mu}{a\ \sigma}\right)^2}. \end{aligned}$$

En résumé, on a pour $a \neq 0$

$$\begin{aligned} f_Y(y) &= \frac{1}{|a|\ \sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y-(b+a\mu)}{a\ \sigma}\right)^2} \\ &\Longrightarrow\ Y \simeq \mathcal{N}\left(a\mu + b, a^2\sigma^2\right). \end{aligned}$$

Par ailleurs

$$\mathbf{E}(Y) = \mathbf{E}(aX + b) = a\mathbf{E}(X) + b = a\ \mu + b$$

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

3.5.6 Lois déduites de la loi normale

Loi du Khi deux

La loi du χ^2 est une loi très classique en statistique. Elle est liée au test du χ^2 qui permet, par exemple, de savoir si un échantillon donné est

en adéquation avec une loi de probabilité définie à priori. Ou bien dans le cadre de la comparaison de proportions ou le test d'indépendance de deux caractères qualitatifs.
Soient $X_1,\ X_2,\ \ldots,\ X_n,\ n$ v. a. normales centrées réduites indépendantes. On appelle χ_2 la v. a. définie par

$$\chi_n^2 = \sum_{i=1}^{n} X_i^2.$$

On dit que que χ_n^2 suit une loi de Pearson à n degrés de liberté noté (n d. d. l.). Sa fonction de densité est donnée par

$$f_X(x) = \begin{cases} C(n)x^{\frac{n}{2}-1}e^{-\frac{x^2}{2}} & \text{si } x \in \mathbb{R}^* \\ 0 & \text{sinon} \end{cases},$$

avec $C(1) = \frac{1}{\sqrt{2\pi}}$. L'espérance de la v. a. du χ_n^2 est $\mathbf{E}(X) = n$ et de variance $\mathbf{Var}(X) = 2n$.
Il existe une table de la loi du χ^2 qui englobe les valeurs de la fonction de répartition, les valeurs de n et le seuil α.

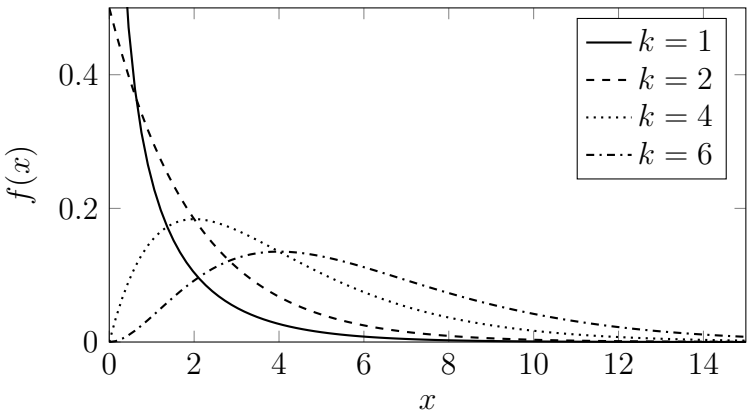


FIGURE 3.10 – Densité de la loi Khi-deux

D'après la figure (3.5.6), on remarque que cette loi est dissymétrique et tend à devenir symétrique pour n assez grand et aussi se rapproche de la répartition normale lorsque $n > 30$.

Loi Student

Soient X une v.a normale standard ($X \simeq \mathcal{N}(0;1)$) et Y une v.a indépendante de X ayant la loi de khi-deux à n degrés de liberté ($Y \simeq \chi^2_n$). La v.a T définie par

$$T = \frac{X}{\sqrt{\frac{Y}{n}}},$$

suit la loi de Student à n degrés de liberté sa densité de probabilité est définie par

$$f(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{2}\right)^{-\frac{n+1}{2}} ; x \in \mathbb{R},$$

et on écrit $T \simeq \mathcal{T}(n)$.

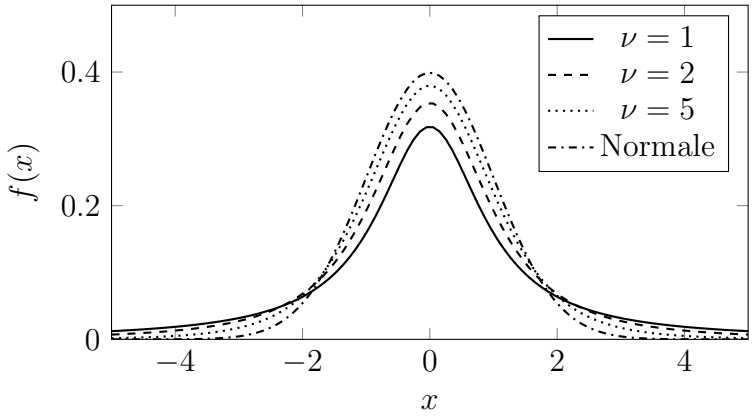


Figure 3.11 – Densité de la loi Student

L'espérance et la variance de T sont $\mathbf{E}(T) = 0$ et $\mathbf{Var}(T) = \frac{n}{n-2}$. Elle est symétrique, plus n est grand et plus sa distribution se confond avec celle de la loi normale standard.

Loi Fisher

Soient X et Y deux v.a indépendantes ayant la loi de khi-deux à respectivement n et m degrés de liberté ($(X \simeq \chi^2_n),\ Y \simeq \chi^2_m$). La v.a F définie par

$$F = \frac{\frac{X}{n}}{\frac{Y}{m}},$$

suit la loi de Fisher-Snedecor à $(n,\ m)$ degrés de liberté, sa densité de probabilité est définie par

$$f(x) = \begin{cases} n^{\frac{n}{2}} m^{\frac{m}{2}} \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \frac{x^{\frac{n}{2}-1}}{(nx+m)^{\frac{n+m}{2}}} & \text{si } x \in \mathbb{R}_+ \\ 0 & \text{sinon} \end{cases},$$

et on écrit $F \simeq \mathcal{F}(n,\ m)$.

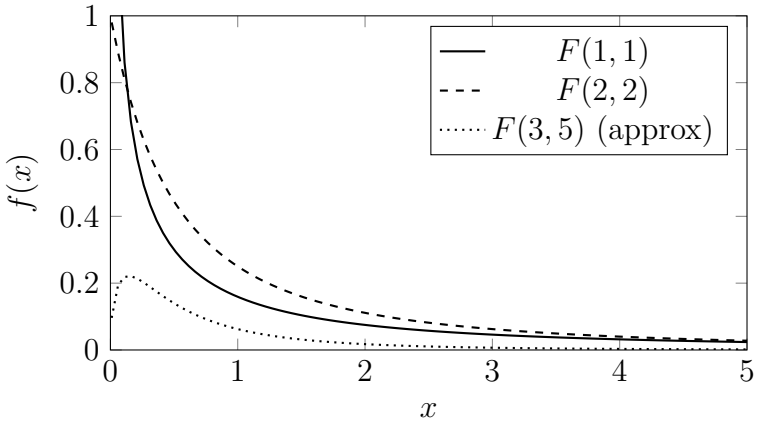


Figure : Densité de la loi de Fisher

L'espérance et la variance de F sont $\mathbf{E}(F) = \frac{m}{m-2}$, si $m > 2$ et $\mathbf{Var}(F) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}$, si $m > 4$.

3.6 Approximation par une loi normale

Dans la suite, on présente l'approximation de loi discrète vers la loi continue normale et cela lorsque la taille de l'échantillon n est assez grand.

3.6.1 Approximation normale d'une répartition binomiale

Soit X une v.a qui suit la loi binomiale, $\mathcal{B}(n; p)$. On a, $\mathbf{E}(X) = np$ et $\mathbf{Var}(X) = np(1-p)$. Et soit la v.a $Z = \frac{X-np}{\sqrt{np(1-p)}}$. On a $\mathbf{E}(Z) = 0$ et $\mathbf{Var}(Z) = 1$.
En pratique, la loi $\mathcal{B}(n; p)$ peut être approchée par la loi normale $\mathcal{N}(np; np(p-1))$ si $n > 30,\ np > 15,\ et\ np(1-p) > 5$. Le théorème suivant le justifie.

Théorème 7 *Théorème de Moivre-Laplace*

Soit X une v.a telle que $X \simeq \mathcal{B}(n,p)$; alors pour tout $a<b$ $(a,b\in\mathbb{R})$ on peut écrire

$$\mathbf{P}\left(a \leq \frac{X-np}{\sqrt{np\ (1-p)}} \leq b\right) \underset{n\longrightarrow+\infty}{\longrightarrow} F_Z(b) - F_Z(a).$$

En règle générale, cette approximation est tout à fait satisfaisante dès que $n\ p\,(1-p) > 10$.

Exemple 4 *On jette une pièce de monnaie 40 fois. Soit la v.a X qui représente le nombre de piles obtenus. Calculer $\mathbf{P}(X=20)$ par approximation normale puis comparer le résultat avec la valeur exacte.*
Il est clair que la v.a $X \simeq \mathcal{B}\left(40,\frac{1}{2}\right)$, donc

$$\mathbf{P}(X=20) = C_{40}^{20}\left(\frac{1}{2}\right)^{20}\left(\frac{1}{2}\right)^{40-20} = 0.1254 \textit{(la valeur exacte)}.$$

On recalcule cette probabilité par approximation normale

$$\begin{aligned}
\mathbf{P}(X=20) &= \mathbf{P}(19.5<X<20.5)\\
&= \mathbf{P}\left(\frac{19.5-20}{\sqrt{10}} < \frac{X-20}{\sqrt{10}} < \frac{20.5-20}{\sqrt{10}}\right)\\
&= \mathbf{P}\left(-0.16 < \frac{X-20}{\sqrt{10}} < 0.16\right)\\
&= F_Z(0.16) - F_Z(-0.16).
\end{aligned}$$

Comme $Z = \frac{X-20}{\sqrt{10}} \simeq \mathcal{N}(0,1)$, on a

$\mathbf{P}(X=20) = 2\ F_Z(0.16) - 1 = 2\,(0.5636) - 1 = 0.1272$*(valeur approximée)*.

Remarquer que lorsque la v.a $Z \simeq \mathcal{N}$ donc $\mathbf{P}(Z=20) = 0$ or $X \simeq \mathcal{B}(n,p)$, d'où la correction par continuité suivante

$$\mathbf{P}(X=20) = \mathbf{P}(19.5<X<20.5).$$

On peut alors remplacer X par Z.

3.6.2 Approximation normale d'une répartition de Poisson

Soit la v.a X qui suit la loi de poisson de paramètre λ, $X \simeq \mathcal{P}(\lambda)$, avec $\mathbf{E}(X) = \mathbf{Var}(X) = \lambda$. Et soit la v.a $Z = \frac{X-\lambda}{\sqrt{\lambda}}$ on a $\mathbf{E}(Z) = 0$ et $\mathbf{Var}(Z) = 1$. En pratique, la loi $\mathcal{P}(\lambda)$ peut être approchée par la loi normale $\mathcal{N}(0,1)$ si $\lambda > 15$.
Soit F_X la fonction de répartition de la $\mathcal{N}(\lambda,\ \lambda)$, et F_Z la fonction de répartition de la loi normale standard $\mathcal{N}(0,1)$ on a

$$F_X(x) = F_Z\left(\frac{x-\lambda}{\sqrt{\lambda}}\right).$$

$$\begin{aligned}\mathbf{P}(X = k) &\simeq \mathbf{P}(k - 0.5 \leq X \leq k + 0.5) = F_X(k + 0.5) - F_X(k - 0.5)\\ &= F_Z\left(\frac{k+0.5-\lambda}{\sqrt{\lambda}}\right) - F_Z\left(\frac{k-0.5-\lambda}{\sqrt{\lambda}}\right),\end{aligned}$$

de même, pour que

$$\sum_{k=0}^{+\infty} \mathbf{P}(X = k) = 1,$$

on a $\mathbf{P}(X = 0) \simeq F_Z\left(\frac{x-\lambda}{\sqrt{\lambda}}\right)$.

3.7 Exercices

Exercice 24 *Soit X une variable aléatoire suivant une loi $\mathcal{U}([a,b])$. pour $k \in \mathbb{N}$, calculer $\mathbf{E}[X^k]$.*

Exercice 25 *Pour $\theta \in [0,1]$, on considère la fonction suivante*

$$f(x) = \begin{cases} \theta e^{2+x} & si\,;\, x \leq -2 \\ 0 & si\,;\, -2 < x < -1 \\ (1-\theta)e^{-1-x} & si\,; x \geq -1 \end{cases} .$$

1. *Démontrer que f est une densité de probabilité.*
2. *Soit X une variable aléatoire dont la loi admet la densité f. Calculer $\mathbf{E}[X]$ et $\mathbf{Var}[X]$.*

Exercice 26 *Soit X une variable aléatoire réelle de loi uniforme sur l'intervalle $[0,4]$. On définit la variable aléatoire Y par $Y = X^2 - 4X + 3$.*

1. *Déterminer f_Y la densité de la loi de Y.*
2. *Montrer qu'on a $\mathbf{P}(Y \leq 0) = \mathbf{P}(Y \geq 0)$.*
3. *Pour $\theta \in \mathbb{R}$, on définit la variable aléatoire Z par*

$$Z = \begin{cases} \theta & si\,; Y \leq 0 \\ Y & sinon \end{cases} .$$

— *Déterminer $\mathbf{E}[Z]$ et $\mathbf{Var}[Z]$.*

Exercice 27 *Rappeler la densité de la loi $\mathcal{N}(\mu, \sigma^2)$. Soit $X \simeq \mathcal{N}(\mu, \sigma^2)$, quelle est la loi de $Y = \frac{X-\mu}{\sigma}$? Calculer $\mathbf{E}[e^{\lambda Y}]$ pour tout $\lambda \in \mathbb{R}$, et en déduire $\mathbf{E}[e^{\lambda X}]$ pour tout $\lambda \in \mathbb{R}$.*

Exercice 28 *Soit X une v.a. réelle de fonction de répartition F_X. Trouver en fonction de F_X les fonctions de répartition de X^2, X^3, $[X]$, (où $[X]$ est la partie entière de X) et $exp(X)$.*

4 Vecteurs aléatoires

Ce chapitre concerne l'étude statistique simultané de plusieurs caractères distincts sur une même population dans le but de retrouver les relations qui peuvent les relier. Par exemple, si un médecin cardiovasculaire suit un malade, il est judicieux de suivre l'évolution de sa tension, sa pression sanguine, son diabète et autres caractéristiques médicaux. On définit ce type de cas de figure, en caractérisant leur lois jointes ainsi que les propriétés induites de l'écriture spécifique dite vecteur aléatoire. On introduit les méthodes d'utilisation théorique, comme l'extraction des lois marginales et des lois conditionnelles.

Dans tout le chapitre, on considère l'écriture $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ ou (X, Y) comme vecteur colonne, pour soulager l'écriture.

4.1 Définitions et propriétés

Soient $X_1,\ \ldots,\ X_n,\ n$ v. a. r. définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$. En pratique, on adopte plus volontiers un point de vue équivalent qui consiste à identifier $\mathbf{X}$ à ses composantes $\mathbf{X} = (X_1,\ X_2,\ \ldots,\ X_n)$. $\mathbf{X}$ est ainsi un n-uplet de variables aléatoires réelles. Réciproquement, la juxtaposition de n v.a.r. quelconques permet d'obtenir un vecteur aléatoire réel noté ($\overrightarrow{v}$. a.).

La loi de $\mathbf{X}$ est la mesure de probabilité $\mathbf{P}_{\mathbf{X}}$ sur $\mathbb{R}^n$ définie, $\forall I = I_1,\ \ldots,\ I_n \subset \mathbb{R}^n$, par

$$\mathbf{P}_X(I) = \mathbf{P}(X \in I) = \mathbf{P}(X_1 \in I_1,\ \ldots,\ X_n \in I_n).$$

Les lois des v. a. r. prises séparément sont dites lois marginales du $\overrightarrow{v}$. a. $\mathbf{X}$.

4.1.1 Vecteur aléatoire discret

$\mathbf{X}$ est un $\overrightarrow{v}$. a. si $\mathbf{X}(\Omega)$ est au plus dénombrable et sa loi de probabilité est déterminée par

$$\forall(x_1,\ \ldots,\ x_n) \in \mathbf{X}(\Omega) : \mathbf{P}[X_1 = x_1,\ \ldots,\ X_n = x_n].$$

Les lois marginales sont toutes discrètes et on a $\mathbf{P}[X_i = x_i]$

$$= \sum_{x_1 \in \mathbf{X}(\Omega)} .. \sum_{x_{i-1} \in \mathbf{X}(\Omega) x_{i+1} \in \mathbf{X}(\Omega)} ..$$

$$\sum_{x_n \in \mathbf{X}(\Omega)} \mathbf{P}[X_1 = x_1, .., X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, .., X_n = x_n].$$

La proposition suivante s'applique aux couples de variables aléatoires. La démonstration est très importante pour l'application des propriétés des probabilités.

Proposition 6 *Soient $(X,\ Y)$ un couple de variables aléatoires discrètes de support $S_{(X,\ Y)} = \{(x_i,\ y_j)/i \in \mathbb{N}^*, j \in \mathbb{N}^*\} = S_X \times S_Y$ les lois marginales $\mathbf{P}_X,\ \mathbf{P}_Y$ (i. e les lois respectives de X et de Y) sont*

$$\forall i \in \mathbb{N}^*; \mathbf{P}_X[\{x_i\}] = \sum_{j \in \mathbb{N}*} \mathbf{P}(X = x_i, Y = y_j),$$

$$\forall j \in \mathbb{N}^*; \mathbf{P}_Y[\{y_j\}] = \sum_{i \in \mathbb{N}*} \mathbf{P}(X = x_i, Y = y_j).$$

Preuve 7 *On a $\forall i \in \mathbb{N}^*$*

$$\mathbf{P}_X[\{x_i\}] = \mathbf{P}[X = x_i] = \mathbf{P}[\{X = x_i\} \cap \Omega]$$

$$= \mathbf{P}[\{X = x_i\} \cap \{Y^{-1}(S_Y)\}] = \mathbf{P}[\{X = x_i\} \cap \{Y(\omega) \in \{y_1, \ldots, y_n\}]$$

$$= \mathbf{P}[\{X = x_i\} \cap \cup_{j \in \mathbb{N}^*} \{Y(\omega) = y_j\}]$$

$$\overset{\text{de l'indépendance des événements}}{=} \sum_{j \in \mathbb{N}^*} \mathbf{P}(X = x_i, Y = y_j).$$

D'une manière analogue, on montre la deuxième formule.

Exemple 6 *Soient $(X,\ Y)$ $\overrightarrow{v}$. a. de loi de probabilité résumé dans le tableau de contingence.*

$X\downarrow \setminus Y\rightarrow$	*0*	*1*	*2*	*3*	*loi de* $X\downarrow$
0	$\frac{2}{21}$	$\frac{1}{21}$	$\frac{2}{21}$	*0*	$\frac{5}{21}$
1	*0*	$\frac{3}{21}$	$\frac{2}{21}$	$\frac{1}{21}$	$\frac{6}{21}$
2	$\frac{2}{21}$	$\frac{1}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{10}{21}$
loi de $Y\rightarrow$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{7}{21}$	$\frac{5}{21}$	$\frac{21}{21}=1$

Par exemple, la loi de X *est donnée par*

X	*0*	*1*	*2*	*somme*
$\mathbf{P}[X=x_i]$	$\frac{5}{21}$	$\frac{6}{21}$	$\frac{10}{21}$	$\frac{21}{21}=1$

La loi de Y *est résumé dans le tableau suivant*

Y	*0*	*1*	*2*	*3*	*somme*
$\mathbf{P}[Y=y_j]$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{7}{21}$	$\frac{5}{21}$	$\frac{21}{21}=1$

On a obtenu

$$\mathbf{P}[X=2]=\sum_{j=1}^{4}\mathbf{P}[X=2,\ Y=y_j]$$

$$=\mathbf{P}[X=2,\ Y=0]+\mathbf{P}[X=2,\ Y=1]+\mathbf{P}[X=2,\ Y=2]+\mathbf{P}[X=2,\ Y=3]=\frac{10}{21}.$$

Exemples théoriques de vecteur aléatoire

Comme dans le cas de v. a., il existe des vecteurs aléatoires de loi usuelle. Par exemple, la loi hypergéométrique multiple et la loi multinomiale.

1. **La loi hypergéométrique multiple**
   Une urne contient $n\geq 2$ couleurs différentes et on considère N_1 de la première couleur,..., N_n de la $n^{\text{ième}}$ couleur. On tire d boules de cette urne et on note X_i le nombre de boules de couleur i tirées. Donc $X_i\simeq\mathcal{H}(N_i,\ N-N_i,d)$ et $N=\sum_{i=1}^{n}N_i$. Pour tout n-uplet $(x_1,\ \ldots,\ x_n)\in\mathbb{N}^n$ tel que $\sum_{i=1}^{n}x_i=d$ et $x_i\leq N_i$, on a
   $$\mathbf{P}(X_1,\ \ldots,\ X_n=x_n)=\frac{C_{N_1}^{x_1}C_{N_2}^{x_2}\ldots C_{N_n}^{x_n}}{C_N^d}.$$

2. **La loi multinomiale**
   Dans le même contexte, si $\forall i=1,\ldots,n:\ \frac{N_i}{N}\underset{N_i\to\infty}{\longrightarrow}p_i$, alors la loi

hypergéométrique multiple converge vers la loi multinomiale de taille d et de $(p_1, \ldots, p_n)$ paramètres définie par

$$\mathbf{P}[X_1 = x_1, \ldots, X_n = x_n] = \frac{d!}{x_1! \ldots x_n!} p_1^{x_1} \ldots p_n^{x_n}.$$

4.1.2 Vecteur aléatoire continu

On dit qu'une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une densité de probabilité si

1. $f \geq 0$
2. et $\int_{\mathbb{R}} \ldots \int_{\mathbb{R}} f_{\mathbf{X}}(x_1, \ldots, x_n) dx_1 \ldots dx_n = 1$.

On dit qu'un $\overrightarrow{v}$. a. $\mathbf{X} = (X_1, \ldots, X_n)$, admet pour densité $f_{\mathbf{X}}$ si $\forall I = I_1 \times \cdots \times I_n$

$$\mathbf{P}[\mathbf{X} \in I] = \int_{I_1} \ldots \int_{I_n} f_{\mathbf{X}}(x_1, \ldots, x_n) dx_1 \ldots dx_n.$$

Pour obtenir la densité marginale de X_i, on a

$$f_{X_i}(x_i) = \underbrace{\int_{\mathbb{R}} \ldots \int_{\mathbb{R}}}_{(n-1) fois} f_{\mathbf{X}}(x_1, \ldots, x_n) dx_1 \ldots dx_{i-1} dx_{i+1} dx_n.$$

La proposition suivante concerne un couple de variables aléatoires. Ce qui correspond à un vecteur aléatoire de dimension deux.

Proposition 7 *Soit $(X,\ Y)$ un couple de v.a. de densité conjointe $f_{(X,\ Y)} : \mathbb{R}^2 \longrightarrow \mathbb{R}$. Alors, les deux v. a. ont une densité marginale notée f_X et f_Y et données par*

$$f_X(x) = \int_{\mathbb{R}} f_{(X,\ Y)}(x,\ y) dy, \ \textit{et}\ f_Y(y) = \int_{\mathbb{R}} f_{(X,\ Y)}(x,\ y) dx.$$

La démonstration se fait à l'aide du théorème de Fubini-Tonelli qui équivaut l'intégration successives.

Exemple 7 *Soit $f_{(X,\ Y)} = \alpha \mathbf{1}_{[0,\ 1] \times [-1,\ 2]}(x,\ y)$.*

1. *Trouver α pour que $f_{(X,\ Y)}$ soit une densité de probabilité.*
2. *Donnez les densités marginales.*

Solution

1. *Pour que $f_{(X,\ Y)}$ soit une densité il faut que*
   - $f_{(X,\ Y)} \geq 0 \Rightarrow \alpha \geq 0$.
   - $\int_{\mathbb{R}^2} f_{(X,\ Y)}(x,\ y) = \alpha \int_0^1 \left(\int_{-1}^2 dy \right) dx = 1$

$$\Rightarrow \alpha \int_{-1}^{2} dy \int_0^1 dx = 1$$

$$\Rightarrow \alpha(y\ |_{-1}^{2})(x\ |_0^1) = \alpha(3)(1) = 1 \Rightarrow \alpha = \frac{1}{3}.$$

   Donc $f_{(X,\ Y)} = \frac{1}{3}\mathbf{1}_{[0,\ 1]\times[-1,\ 2]}(x,\ y)$.

2. *Densité de X*

$$f_X(x) = \int_{\mathbb{R}} f_{(X,\ Y)}(x,\ y)dy$$

$$= \frac{1}{3}\mathbf{1}_{[0,\ 1]}(x) \int_{-1}^{2} dy = \frac{1}{3}\mathbf{1}_{[0,\ 1]}(x)(3) = \mathbf{1}_{[0,\ 1]}(x).$$

   Densité de Y

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,\ Y)}(x,\ y)dx$$

$$= \frac{1}{3}\mathbf{1}_{[-1,\ 2]}(y) \int_0^1 dx = \frac{1}{3}\mathbf{1}_{[-1,\ 2]}(y)(1) = \frac{1}{3}\mathbf{1}_{[-1,\ 2]}(y).$$

Dans la suite, on considère que les couples de variables aléatoires pour soulager l'écriture. Mais tout ce qui se dira est vérifié quelque soit la dimension du vecteur.

4.2 La fonction de répartition

Dans les chapitres précédents, on a défini la fonction de répartition et ses propriétés des v. a. discrète ou continue. Dans la suite on va faire de même pour les vecteurs aléatoires. C'est-à-dire que la loi conjointe des v. a. X et Y est constituée par la fonction de répartition du couple aléatoire $(X,\ Y)$, définie sur $\mathbb{R}^2$, par

$$F_{(X,\ Y)}(x,\ y) = \mathbf{P}_{(X,\ Y)}\left( ]-\infty,\ x] \times]-\infty,\ y[\right)$$

$$= \mathbf{P}\left(X \in]-\infty,\ x],\ Y \in]-\infty,\ y[\right)$$

$$= \mathbf{P}\Big(X \leq x,\ Y \leq y \Big).$$

Propriété 6 *Soit $(X,\ Y)$ un couple de v. a. de fonction de répartition $F_{(X,\ Y)}$, alors*

1. $\forall (x,\ y) \in \mathbb{R}^2,\ 0 \leq F_{(X,\ Y)}(x,\ y) \leq 1.$
2. *$F_{(X,\ Y)}$ est une fonction croissante en chacune des ses variables.*
3. *$F_{(X,\ Y)}$ est continue à droite et admet une limite à gauche pour chacune de ses composantes.*
4. $$\lim_{\substack{x \to +\infty \\ et\ y \to +\infty}} F_{(X,\ Y)}(x,\ y) = 1, \qquad \lim_{\substack{x \to -\infty \\ ou\ y \to -\infty}} F_{(X,\ Y)}(x,\ y) = 0.$$
5. *Si $(X,\ Y)$ un couple de v. a. discrètes alors*
   — $$F_{(X,\ Y)}(x,\ y) = \sum_{x_i \leq x} \sum_{y_j \leq y} \mathbf{P}\Big(X = x_i,\ Y = y_j \Big).$$
   — *Les fonctions de répartitions marginales sont données par*
   $$F_X(x) = \lim_{y \to +\infty} F_{(X,\ Y)}(x,\ y) = \sum_{x_i \leq x} \sum_{y_j \leq +\infty} \mathbf{P}\Big(X = x_i,\ Y = y_j \Big).$$
   $$F_Y(y) = \lim_{x \to +\infty} F_{(X,\ Y)}(x,\ y) = \sum_{x_i \leq +\infty} \sum_{y_j \leq y} \mathbf{P}\Big(X = x_i,\ Y = y_j \Big).$$
6. *Si $(X,\ Y)$ un couple de v. a. continues alors*
   — $$F_{(X,\ Y)}(x,\ y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{(X,\ Y)}(u,\ v) du dv.$$
   — *Les fonctions de répartitions marginales sont données par*
   $$F_X(x) = \lim_{y \to +\infty} F_{(X,\ Y)}(x,\ y) = \int_{-\infty}^{x} \int_{-\infty}^{+\infty} f_{(X,\ Y)}(u,\ v) du dv.$$
   $$F_Y(y) = \lim_{x \to +\infty} F_{(X,\ Y)}(x,\ y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{y} f_{(X,\ Y)}(u,\ v) du dv.$$

Proposition 8 *Si $F_{(X,\ Y)}$ est de classe $\mathcal{C}^1(\mathbb{R}^2)$ alors le couple aléatoire $(X,\ Y)$ admet une densité*

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

Exemple 8 *Soit $(X,\ Y)$ un vecteur aléatoire de fonction de répartition $F_{(X,\ Y)}$ donnée par*

$$F_{(X,\ Y)}(x,\ y) = \begin{cases} e^{-\alpha x - \beta y} - e^{-\alpha x} - e^{-\beta y} + 1 & si \quad x,\ y > 0 \\ 0\ sinon \end{cases},$$

où $\alpha,\ \beta$ sont des constantes positives. On détermine la densité jointe. Comme $F_{(X,\ Y)}$ est de classe $\mathcal{C}^1(\mathbb{R}^2)$ alors

$$f_{(X,\ Y)}(x,\ y) = \frac{\partial^2 F_{(X,\ Y)}(x,\ y)}{\partial x \partial y}$$

$$= \frac{\partial^2}{\partial x \partial y}\left\{e^{-\alpha x - \beta y} - e^{-\alpha x} - e^{-\beta y} + 1\right\}$$

$$= \frac{\partial^2}{\partial y}\left\{-\alpha e^{-\alpha x - \beta y} + \alpha e^{-\alpha x}\right\}$$

$$= \alpha\beta e^{-\alpha x - \beta y}.$$

Donc

$$f_{(X,\ Y)}(x,\ y) = \begin{cases} \alpha\beta e^{-\alpha x - \beta y} & si \quad x,\ y > 0 \\ 0\ sinon \end{cases}.$$

4.3 Les moments

Dans cette partie, on considère $\|\|$ la norme dans $\mathbb{R}^2$. Le théorème suivant représente un développement du théorème 1 et le théorème 3 dans le cas de variable aléatoire.

Théorème 8 *Théorème de transfert*
Soit $\mathbf{X}$ un $\overrightarrow{v}$. a. de $\mathbb{R}^n$ et $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction borélienne.

1. *Si $\mathbf{X}$ est discret et si la série $\sum_{x \in S_{\mathbf{X}}} \|\varphi(x)\| \mathbf{P}[\mathbf{X} = x]$ converge*

$$\Rightarrow \mathbf{E}\left[\varphi(\mathbf{X})\right] = \sum_{x \in S_{\mathbf{X}}} \varphi(x) \mathbf{P}[\mathbf{X} = x].$$

2. Si $\mathbf{X}$ *est continu de densité* $f_{\mathbf{X}}$ *et si* $\forall x \in \mathbb{R}^n : \|\varphi(x)\| f_{\mathbf{X}}$ *intégrable*

$$\Rightarrow \mathbf{E}\left[\varphi(\mathbf{X})\right] = \int_{S_{\mathbf{X}}} \varphi(x) f_{\mathbf{X}}(x) dx.$$

Si on considère $\varphi(x) = x,\ \varphi(x) = x^2$ ou bien pour deux vecteurs, $\varphi(x, y) = xy$ on obtient ce qui suit. Soit $\mathbf{X} = (X_1, \ldots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire (discret ou continu).

Définition 7 *L'espérance d'un vecteur aléatoire* $\mathbf{X}$ *de dimension* n *est le vecteur de* $\mathbb{R}^n$ *constitue des espérances de chacune des coordonnées i.e.*

$$\mathbf{E}(X) = \left(\mathbf{E}(X_1), \ldots, \mathbf{E}(X_n)\right).$$

Proposition 9 *(Linéarité de l'espérance) Soit* $B \in \mathbb{R}^m$ *et* $A \in \mathbb{R}^{m \times n}$. *L'espérance du vecteur aléatoire* $Y = AX + B \in \mathbb{R}^m$ *est*

$$\mathbf{E}(Y) = A\mathbf{E}(X) + B.$$

Cas particulier $\mathbf{E}(X_1 + X_2 + \cdots + X_n) = \mathbf{E}(X_1) + \mathbf{E}(X_2) + \cdots + \mathbf{E}(X_n)$.

Soit $\mathbf{X} = (X_1, \ldots, X_n) \in \mathbb{R}^n$ un vecteur aléatoire (discret ou continu).
La covariance entre X_i et X_j

$$\mathbf{Cov}(X_i,\ X_j) = \mathbf{E}\left((X_i - \mathbf{E}(X_i))(X_j - \mathbf{E}(X_j))\right) = \mathbf{E}(X_i X_j) - \mathbf{E}(X_i)\mathbf{E}(X_j).$$

Le coefficient de corrélation linéaire entre X_i et X_j

$$\rho(X_i, X_j) = \frac{\mathbf{Cov}(X_i, X_j)}{\sqrt{\mathbf{Var}(Xi)}\sqrt{\mathbf{Var}(Xj)}} \in [-1,\ 1].$$

Définition 8 *Soit* $X = (X_1; \ldots; X_n)$ *un vecteur aléatoire de dimension* n. *On appelle matrice de variance-covariance de* X, *la matrice carrée de taille* $n \times n$, *notée* Σ_X, *dont les coefficients* $(\Sigma_{i;j})_{1 \leq i;j \leq n}$ *sont donnés par*

$$\Sigma_{i;j} = \mathbf{Cov}(X_i; X_j).$$

En notation matricielle, la matrice de variance-covariance d'un vecteur aléatoire $\mathbf{X}$ est

$$\Sigma_{\mathbf{X}} = \mathbf{E}((\mathbf{X} - \mathbf{E}(\mathbf{X}))(\mathbf{X} - \mathbf{E}(\mathbf{X}))^t).$$

On appelle matrice de variance-covariance du couple $(X,\ Y)$ la matrice carrée d'ordre 2 symétrique définie par

$$\Sigma_{(X,\ Y)} = \mathbf{E}[(X - \mathbf{E}(X),\ Y - \mathbf{E}(Y)),\ (X - \mathbf{E}(X),\ Y - \mathbf{E}(Y))^t]$$

$$= \begin{pmatrix} \mathbf{Var}(X) & \mathbf{Cov}(X;Y) \\ \mathbf{Cov}(X;Y) & \mathbf{Var}(Y) \end{pmatrix},$$

ses termes diagonaux sont les variances des composantes du couple.

Remarque 9

- *La diagonale de Σ_X est le vecteur $(\mathbf{Var}(X_1); \ldots; \mathbf{Var}(X_n))$.*
- *Si les coordonnées de $\mathbf{X}$ sont indépendantes, alors $\Sigma_{\mathbf{X}}$ est une matrice diagonale.*

Proposition 10 *Soient $(X;Y) \in \mathbb{R}^2$ un couple de v.a.r. (discret ou continu). Alors,*

$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y) + 2\mathbf{Cov}(X;Y).$$

La démonstration est laissée aux étudiants.

4.4 Loi conditionnelle et indépendance

A partir de la loi jointe d'un vecteur aléatoire $\mathbf{X}$ (discret ou continu), on peut aboutir à de nouvelles lois en indiquant une condition sur l'une ou plusieurs composantes de $\mathbf{X}$. Ce cas de figure est dit lois conditionnelles. Dans la suite, on considère le cas d'un couple de v. a. mais ça peut se généraliser aux vecteurs de taille supérieure.

4.4.1 Lois conditionnelles

Soient X et Y deux v. a. (discrètes ou continues) sur le même espace $(\Omega, \mathcal{A}, \mathbf{P})$.

1. Si $\mathbf{X} = (X,\ Y)$ est discret de loi de probabilité $\mathbf{P}[X = x_i,\ Y = y_j] = p_{ij}$ et de lois marginales $\mathbf{P}[X = x_i] = p_i,\ \mathbf{P}[Y = y_j] = p_j,\ \forall x_i \in$

$S_X, y_j \in S_Y$, alors la loi conditionnelle de X sachant $Y = y_j$ est donnée pour $p_j \neq 0$ par

$$\mathbf{P}[X = x_i | Y = y_j] = \frac{\mathbf{P}[X = x_i,\ Y = y_j]}{\mathbf{P}[Y = y_j]} = \frac{p_{ij}}{p_j}.$$

Et de même la loi conditionnelle de Y sachant $X = x_i$ est donnée pour $p_i \neq 0$ par

$$\mathbf{P}[Y = y_j | X = x_i] = \frac{\mathbf{P}[X = x_i,\ Y = y_j]}{\mathbf{P}[X = x_i]} = \frac{p_{ij}}{p_i}.$$

2. Si $\mathbf{X} = (X,\ Y)$ est continu de densité de probabilité $f_{(X,\ Y)}$ et de densités marginales $f_X,\ f_Y,\ \forall x,\ y \in S_{(X,\ Y)}$, alors la densité conditionnelle de X sachant $Y = y$ est donnée pour $f_Y(y) \neq 0$ par

$$f_{[X|Y=y]}(x) = \frac{f_{(X,\ Y)}(x, y)}{f_Y(y)}.$$

Et de même la densité conditionnelle de Y sachant $X = x$ est donnée pour $f_X(x) \neq 0$ par

$$f_{[Y|X=x]}(y) = \frac{f_{(X,\ Y)}(x, y)}{f_X(x)}.$$

Exemple 9 *En reprenant l'exemple (6) dans le cas discret, on a par exemple la loi de X sachant que $Y = 1$*

$$\mathbf{P}(X = 0 | Y = 1) = \frac{\mathbf{P}(X = 0,\ Y = 1)}{\mathbf{P}(Y = 1)} = \frac{1/21}{5/21} = \frac{1}{5}.$$

$$\mathbf{P}(X = 1 | Y = 1) = \frac{\mathbf{P}(X = 1,\ Y = 1)}{\mathbf{P}(Y = 1)} = \frac{3/21}{5/21} = \frac{3}{5}.$$

$$\mathbf{P}(X = 2 | Y = 1) = \frac{\mathbf{P}(X = 2,\ Y = 1)}{\mathbf{P}(Y = 1)} = \frac{1/21}{5/21} = \frac{1}{5}.$$

Et on a $\sum_{i=1}^{3} \mathbf{P}(X = x_i | Y = 1) = 1$.

Exemple 10 *En reprenant l'exemple (7) dans le cas continu, on a par exemple la densité de Y sachant que $X = x$*

$$f_{[Y|X=x]}(y) = \frac{f_{(X,\ Y)}(x, y)}{f_X(x)} = \frac{\frac{1}{3}\mathbf{1}_{[0,\ 1]\times[-1,\ 2]}(x,\ y)}{\mathbf{1}_{[0,\ 1]}(x)} = \frac{1}{3}\mathbf{1}_{[-1,\ 2]}(y).$$

Aussi $\int_{\mathbb{R}} \frac{1}{3}\mathbf{1}_{[-1,\ 2]}(y)dy = \int_{-1}^{2} \frac{1}{3}dy = 1$.

Ces loi conditionnelles peuvent ne pas exister si l'indépendance entre les composantes du vecteurs est prouvée.

4.4.2 Rappel d'indépendance de variables aléatoires

Soient X et Y deux v. a. (discrètes ou continues) sur le même espace $(\Omega, \mathcal{A}, \mathbf{P})$. X et Y sont indépendantes si pour tout couple d'intervalles I et J de $\mathbb{R}$, on a

$$\mathbf{P}(X \in I,\ Y \in J) = P(X \in I) \times P(Y \in J).$$

Un résultat direct de cette définition est la proposition suivante.

Proposition 11 *Soit $(X;Y)$ un couple aléatoire (discret ou continu). Si X et Y sont indépendantes, alors*

1. *Si $(X;Y)$ discret, on a*

$$\mathbf{P}[X = x_i,\ Y = y_j] = \mathbf{P}[X = x_i] \times \mathbf{P}[Y = y_j],\ \forall x_i \in S_X,\ y_j \in S_Y.$$

2. *Si $(X;Y)$ continu, on a*

$$f_{(X,\ Y)}(x, y) = f_X(x) f_Y(y),\ \forall x \in S_X,\ y \in S_Y.$$

3. *Si $(X;Y)$ discret ou continu*

$$F_{(X,\ Y)}(x, y) = F_X(x) \times F_Y(y).$$

4. *Si $(X;Y)$ discret ou continu*

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

5. *Si $(X;Y)$ discret ou continu*

$$\mathbf{Cov}(X,\ Y) = 0,$$

mais la réciproque n'est pas toujours vraie.

6. *Si X et Y sont indépendantes, alors*

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

7. *Pour a et b deux réels*

$$\mathbf{Cov}(aX_1 + bX_2,\ Y) = a\ \mathbf{Cov}(X_1,\ Y) + b\ \mathbf{Cov}(X_2,\ Y).$$

La démonstration de cette proposition se fera comme exercice pour les étudiants.

Proposition 12 *Soit $(X;Y)$ un couple aléatoire (discret ou continu). Si* $\mathbf{Cov}(X,\ Y) = 0$, *alors X et Y ne sont pas forcément indépendantes.*

La preuve de cette proposition est donnée par un contre exemple.

4.5 Calcul de loi

Soit $\mathbf{X} = (X_1,\ \ldots,\ X_n)$ un vecteur aléatoire de loi $\mathbf{P}_{\mathbf{X}}$ et de fonction de répartition $F_{\mathbf{X}}$. On considère $\varphi : \mathbb{R}^n \to \mathbb{R}^p$ une fonction mesurable. On se propose la problème de calculer la loi de $Y = \varphi(X)$.

Cas discret

Pour une loi discrète, l'approche s'effectue généralement en deux temps : détermination du support S_Y de Y, puis, on calcul des probabilités $\mathbf{P}(Y = y) = \mathbf{P}(\varphi(X) = y)$ pour tout $y = (y_1,\ \ldots,\ y_p) \in S_Y$.

Cas absolument continue

Utilisation de la fonction de répartition est une idée qui consiste à exprimer la fonction de répartition F_Y en fonction de F_X comme suit

$$F_Y(y) = \mathbf{P}\left(\varphi(\mathbf{X}) \in]-\infty; y_1] \times \cdots \times]-\infty; y_p]\right)$$

$$= \int \mathbf{1}_{\varphi(x)\in]-\infty;y_1]\times\cdots\times]-\infty;y_p]} d\mathbf{P}_{\mathbf{X}}(x_1;\ldots;x_n).$$

Une fois cette intégrale calculée, on obtient la densité de Y en dérivant F_Y par rapport à ses p variables, donnée par

$$f_Y(y) = \frac{\partial F_Y}{\partial y_1 \ldots \partial y_p}(y_1,\ \ldots,\ y_p).$$

Le calcul de F_Y est souvent long. On préfère utiliser les méthodes suivantes. Utilisation de la mesure image où on se place ici dans le cas où S_X est un ouvert U inclus dans R^n (la densité s'écrit $f_X(x)\mathbf{1}_U(x)$) et on suppose que

$\varphi : U \to V$ est une $\mathcal{C}^1$-difféomorphisme (différentiable, bijective et dont les dérivées partielles sont continues). On a alors le théorème de changement de variable suivant.

Théorème 9 *Sous les hypothèses ci-dessus, on a pour toute fonction f intégrable*

$$\int_V f(y)d\lambda(y) = \int_U f(\varphi(x))|det(J_{\varphi(x)})|d\lambda(x),$$

et

$$\int_U f(x)d\lambda(x) = \int_V f(\varphi^{-1}(y))|det(J_{\varphi^{-1}(y)})|d\lambda(y),$$

où $J_{\varphi(x)}$ désigne la matrice Jacobienne au point x

$$J_{\varphi(x)} = \begin{pmatrix} \frac{\partial\varphi_1}{\partial x_1}(x) & \cdots & \frac{\partial\varphi_1}{\partial x_1}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial\varphi_n}{\partial x_n}(x) & \cdots & \frac{\partial\varphi_n}{\partial x_n}(x) \end{pmatrix}.$$

On déduit alors la densité de Y.

Théorème 10 *$\mathbf{P}_Y$ est la mesure image de $\mathbf{P}_X$ par ($\mathbf{P}_Y = \mathbf{P}_X \circ \varphi^{-1}$). On a donc*

$$f_Y(y) = f_X(\varphi^{-1}(y))|det(J_{\varphi^{-1}y})|\mathbf{1}_V(y).$$

Exemple 11 *Soit $X = (X_1; X_2)$ de loi uniforme sur le carré $D =]0;1[^2$. On cherche la loi de $X_1 + X_2$. On pose $Y = (X_1 + X_2;\ X_2)$ et*

$$\begin{array}{rcl} \varphi : U & \to & V = \{(y_1,\ y_2),\ 0 < y_2 < 1 \text{ et } y_2 < y_1 < 1 + y_2\} \\ (x_1,\ x_2) & \mapsto & (y_1 = x_1 + x_2,\ y_2 = x_2) \end{array}.$$

D'après le théorème 10, la densité de Y est donnée par

$$f_Y(y) = \mathbf{1}_U(\varphi^{-1}(y))|det(J_{\varphi^{-1}(y)})|\mathbf{1}_V(y) = 1_{0<y_1-y_2<1}\mathbf{1}_{0<y_2<1}.$$

La densité de $Y_1 = X_1 + X_2$ s'obtient en intégrant f_Y par rapport à y_2

$$f_{Y_1}(y_1) = \int_{\mathbb{R}} f_Y(y)dy_2 = y_1\mathbf{1}_{0<y_1<1} + (2 - y_1)\mathbf{1}_{1<y_1<2}.$$

Utilisation de la fonction muette est une approche identique à celle effectuée pour les v.a.r. Il s'agit de trouver une fonction g mesurable telle que pour toute fonction h mesurable on ait

$$\mathbf{E}[h(\varphi(X))] = \int_U h(\varphi(x)) f_X(x) d\lambda(x) = \cdots = \int_V h(y) g(y) d\lambda(y).$$

Par identification, une telle fonction g est une densité de Y. Cette fonction s'obtient généralement par un changement de variable.

Exemple 12 *On reprend l'exemple précédent. On effectue le changement de variable*

$$\begin{array}{ccc} h : \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (y_1,\ y_2) & \mapsto & (y_1 - y_2,\ y_2) \end{array}.$$

On obtient

$$\mathbf{E}[h(X_1 + X_2)] = \int_{\mathbb{R}^2} h(x_1 + x_2) \mathbf{1}_U(x_1; x_2) dx_1 dx_2$$

$$= \int_{\mathbb{R}^2} h(y_1) \mathbf{1}_U(y_1 - y_2; y_2) dy_1 dy_2$$

$$= \int_{\mathbb{R}} h(y_1) \int_{\mathbb{R}} \mathbf{1}_U(y_1 - y_2; y_2) dy_2 dy_1.$$

Il vient par identification

$$f_{X_1+X_2}(y_1) = \int_{\mathbb{R}} \mathbf{1} U(y_1 - y_2; y_2) dy_2.$$

Produit de convolution permet de déterminer la densité de la somme de deux variables aléatoires réelles indépendantes.

Définition 9 *Soient f et g deux fonction intégrables. On appelle convolution de f et g la fonction $f * g$ définie par*

$$(f * g)(x) = \int f(x - t) g(t) d\lambda(t).$$

Proposition 13 *Soient f, g et h trois fonctions intégrables. Alors*

$$f * g = g * f \text{ et } (f * g) * h = f * (g * h) \lambda - pp.$$

Théorème 11 *Soient X et Y deux v.a.r. indépendantes de densité f_X et f_Y. Alors $f_X * f_Y$ est la densité de $X_1 + X_2$.*

On peut recalculer la densité de $X_1 + X_2$ de l'exemple précédent en faisant le produit de convolution entre les deux fonctions indicatrices sur $]0; 1[$.

La suite de ce cours concerne la loi normale et son utilisation dans la construction des couples mais peut se généraliser aux vecteurs aléatoires.

4.6 Vecteur Gaussien

4.6.1 Couple Gaussien

Soit $(X,\ Y)$ un couple de v.a de $\mathbb{R}^2$ ayant pour espérance mathématique $(\mu_X,\ \mu_Y)$ et pour matrice de variance-covariance $\Sigma_{(X,\ Y)} = \begin{pmatrix} \mathbf{Var}(X) & \mathbf{Cov}(X;Y) \\ \mathbf{Cov}(X;Y) & \mathbf{Var}(Y) \end{pmatrix}$.
Le couple $(X,\ Y)$ est un couple gaussien, si elle admet pour densité

$$f_{(X,\ Y)}(x;y) = \frac{1}{2\pi\sqrt{det\Sigma}} \exp\left[-\frac{1}{2}(x-\mu_X,\ y-\mu_Y)\Sigma^{-1}(X,\ Y)(x-\mu_X,\ y-\mu_Y)^t\right].$$

Les lois marginales des v.a X et Y sont données par la densité de la normale comme suit

$$f_X(x) = \frac{1}{\sqrt{2\pi\mathbf{Var}(X)}} \exp\left[-\frac{(x-\mu_X)^2}{2\mathbf{Var}(X)}\right],$$

et

$$f_Y(y) = \frac{1}{\sqrt{2\pi\mathbf{Var}(Y)}} \exp\left[-\frac{(y-\mu_Y)^2}{2\mathbf{Var}(Y)}\right].$$

4.6.2 Vecteur Gaussien

La définition d'un vecteur aléatoire gaussien n'est pas très simple.

Définition 10 *On dit que le vecteur aléatoire $\mathbf{X} = (X_1,\ \ldots,\ X_n)$ est un vecteur gaussien si toute combinaison linéaire de ses variables marginales $a_1X_1 + \cdots + a_nX_n$ est une v.a.r. gaussienne.*

Proposition 14 *La loi d'un vecteur aléatoire gaussien $\mathbf{X} = (X_1,\ \ldots,\ X_n)$ est entièrement déterminée par la donnée de son espérance $\mu_{\mathbf{X}} \in \mathbb{R}^n$ et de*

sa matrice de variance-covariance $\Sigma_{\mathbf{X}}$ *(de dimension* $n \times n$*). On note alors* $\mathbf{X} \simeq \mathcal{N}(\mu_{\mathbf{X}},\ \Sigma_{\mathbf{X}})$.

Proposition 15 *(Vecteur gaussien* $\Rightarrow$ *Composantes gaussiennes)*
Si le vecteur $\mathbf{X} = (X_1,\ \ldots,\ X_n)$ *est gaussien, alors chaque variable aléatoire* $X_i,\ i = 1; \ldots; n$ *est gaussienne.*

Le réciproque est fausse.

Exemple 13 *Soit* $X \simeq \mathcal{N}(0; 1)$ *et* ε *une variable aléatoire de Rademacher indépendante de* X*, c'est-à-dire,* $\mathbf{P}(\varepsilon = 1) = \mathbf{P}(\varepsilon = -1) = \frac{1}{2}$*. On considère la variable* $Y = \varepsilon X$ *et le vecteur* $T = (X,\ Y)$*. On a alors*

$$F_Y(u) = \mathbf{P}(Y \leq u)$$

$$= \mathbf{P}(\varepsilon = 1) + \mathbf{P}(\varepsilon X \leq u | \varepsilon = -1)\mathbf{P}(\varepsilon = -1)$$

$$= \frac{1}{2}[\mathbf{P}(X \leq u) + \mathbf{P}(X \geq -u) = \mathbf{P}(X \leq u) = F_X(u).$$

Y *est donc une v.a.r. gaussienne centrée réduite. Les composantes du vecteur* $(X,\ Y)$ *suivent ainsi des lois gaussiennes alors que* $(X,\ Y)$ *n'est pas un vecteur gaussien. En effet, il suffit de voir que* $Z = X + Y$ *n'est pas une v.a.r. gaussienne*

$$\mathbf{P}(Z = 0) = \mathbf{P}((1 + \varepsilon)X = 0) = \mathbf{P}(\varepsilon = -1) = \frac{1}{2}.$$

Or si Z *est une v.a.r. gaussienne on a* $\mathbf{P}(Z = 0) = 0$ *(ou 1 dans le cas d'une v.a.r. centrée dégénérée).*

La réciproque de la proposition 15 est en revanche vraie si les marginales sont indépendantes.

Proposition 16 *(Composantes gaussiennes indépendantes* $\Rightarrow$ *Vecteur gaussien)*
Soit $X_1,\ \ldots,\ X_n$*,* n *variables aléatoires indépendantes. Le vecteur* $\mathbf{X} = (X_1,\ \ldots,\ X_n)$ *est gaussien si et seulement si pour tout* $i \in \{1; \ldots; n\}$ *la variable aléatoire* X_i *est gaussienne.*

Proposition 17 *1.* $\mathbf{X} = (X_1, \ldots, X_n) \simeq \mathcal{N}(\mu_{\mathbf{X}}, \Sigma_{\mathbf{X}})$ *si et seulement si* $\mathbf{X}$ *peut se décomposer sous la forme* $\mathbf{X} = AY + \mu_{\mathbf{X}}$, *avec* $Y \simeq \mathcal{N}(0; I_n)$, A *matrice* $d \times d$ *satisfaisant* $AA' = \Sigma_{\mathbf{X}}$ *et* $rang(A) = rang(\Sigma_{\mathbf{X}})$.

2. Si $\mathbf{X} = (X_1, \ldots, X_n) \simeq \mathcal{N}(\mu_{\mathbf{X}}, \Sigma_{\mathbf{X}})$, A *est une matrice* $d \times n$ *et* $b \in \mathbb{R}^d$, *alors* $A\mathbf{X} + b \simeq \mathcal{N}(A\mu_{\mathbf{X}} + b, A\Sigma_{\mathbf{X}}A')$ *En particulier, si* $\Sigma_{\mathbf{X}}$ *est inversible le vecteur* $\Sigma_{\mathbf{X}}^{-1/2}(\mathbf{X} - \mu_{\mathbf{X}}$ *est gaussien standard.*

On a vu plus haut que dans un cadre général l'indépendance implique la non corrélation mais que la réciproque est fausse. La proposition suivante montre que cette réciproque est valable pour les vecteurs gaussiens.

Proposition 18 *(Vecteur gaussien : Indépendance ↔ Décorrélation)*

1. *Soit* $\mathbf{X} = (X_1, \ldots, X_n)$ *un vecteur gaussien de loi* $\mathcal{N}(\mu_{\mathbf{X}}, \Sigma_{\mathbf{X}})$. *Les variables aléatoires* $X_1, \ldots, X_n$ *sont indépendantes si et seulement si elles sont non corrélées, c'est-à-dire si et seulement si la matrice de variance-covariance* $\Sigma_{\mathbf{X}}$ *est diagonale.*
2. *Soit* $Z = (X_1; \ldots; X_n; Y_1; \ldots; Y_p)$ *un vecteur gaussien. Alors, les vecteurs* $(X_1, \ldots, X_n)$ *et* $(Y_1; \ldots; Y_p)$ *sont indépendants si et seulement si ils sont non corrélés.*

Proposition 19 *(Densité)*
Soit $\mathbf{X} = (X_1, \ldots, X_n) \simeq \mathcal{N}(\mu_{\mathbf{X}}, \Sigma_{\mathbf{X}})$, $\mathbf{X}$ *admet une densité* f_X *si et seulement si* $det(\Sigma_{\mathbf{X}}) \neq 0$. *On a alors*

$$f_{\mathbf{X}}(x) = \frac{1}{(\sqrt{2\pi})^n \sqrt{det(\Sigma_{\mathbf{X}})}} \exp\left(\frac{1}{2}(x - \mu_{\mathbf{X}})'\Sigma_{\mathbf{X}}^{-1}(x - \mu_{\mathbf{X}})\right),$$

où $x = (x_1; \ldots; x_n)$.

Preuve 8 *Comme* $\Sigma_{\mathbf{X}}$ *est inversible, le vecteur* $Y = \Sigma_{\mathbf{X}}^{-1}(x - \mu_{\mathbf{X}}) \simeq \mathcal{N}(0; I_n)$. *Sa densité est donnée par*

$$f_Y(y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_i^2}{2}\right) = \frac{1}{(\sqrt{2\pi})^n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right).$$

Le changement de variable $\varphi : \mathbb{R}^n \to \mathbb{R}^n$, *pour* x *on a* $y = \Sigma_{\mathbf{X}}^{-1}(x - \mu_{\mathbf{X}})$ *nous donne la densité de* $\mathbf{X}$ *par*

$$f_{\mathbf{X}}(x) = f_Y\left(\Sigma_{\mathbf{X}}^{-1}(x - \mu_{\mathbf{X}})|detJ_{\varphi(x)}|\right).$$

φ *étant une application affine, sa matrice jacobienne vaut* $\Sigma_{\mathbf{X}}^{-1/2}$, *d'où le résultat.*

Définition 11 *Une vecteur gaussien* X *dont la matrice de variance-covariance a un déterminant nul est dit dégénéré.*

4.7 Fonction caractéristique d'un couple

Dans cette partie, on notera toujours $\mathbf{X} = (X_1,\ X_2)$ un couple de variables aléatoires et $T = (t_1,\ t_2)$ un couple de $\mathbb{R}^2$.

Définition 12 *Soit* $\mathbf{X} = (X_1,\ X_2)$ *un couple de variables aléatoires. On appelle fonction caractéristique du couple* $(X_1,\ X_2)$ *la fonction* φ *définie sur* $\mathbb{R}^2$, *à valeurs dans* $\mathbb{C}$, *par*

$$\varphi(T) = \varphi_{\mathbf{X}}(T) = \mathbf{E}(e^{i<T,\mathbf{X}>}),\ \forall T \in \mathbb{R}^2,$$

où $< T, \mathbf{X} >$ *est le produit scalaire de* $\mathbb{R}^2$ *défini par*

$$< T, \mathbf{X} >= t_1X_1 + t_2X_2.$$

Par exemple, si le couple $\mathbf{X}$ *admet une densité* $f_{\mathbf{X}}$, *en utilisant le théorème de transfert (8), pour tout* $T \in \mathbb{R}^2$

$$\varphi_{\mathbf{X}}(T) = \mathbf{E}(e^{i<T,\mathbf{X}>}) = \int_{\mathbb{R}^2} \exp\left[i(t_1x_1 + t_2x_2)\right] f_{(X_1,\ X_2)}(x_1, x_2)dx_1dx_2.$$

On peut vérifier, comme pour le cas unidimensionnel, que la fonction caractéristique d'un couple existe toujours. On retrouve les propriétés élémentaires des fonctions caractéristiques.

Proposition 20 *Soit* φ *la fonction caractéristique d'un couple de variables aléatoires, les propriétés suivantes sont toujours vraies*

1. φ *est continue sur* $\mathbb{R}^2$.
2. $\varphi(0;0) = 1$.
3. $\varphi(-T) = \overline{\varphi(T)}$ *pour tout* $T \in \mathbb{R}^2$.
4. $|\varphi(T)| \leq 1$ *pour tout* $T \in \mathbb{R}^2$.

Corollaire 1 *Soit* φ *la fonction caractéristique d'un couple* $(X_1,\ X_2)$. *Alors, pour tout* $(t_1; t_2) \in \mathbb{R}^2$,

$$\varphi_{X_1}(t_1) = \varphi(t_1; 0) \text{ et } \varphi_{X_2}(t_2) = \varphi(0; t_2).$$

Proposition 21 *Soit* $\mathbf{X} = (X_1,\ X_2)$ *un couple de variables aléatoires de fonction caractéristique* φ. *Si* φ *est intégrable sur* $\mathbb{R}^2$, *c'est-à-dire si*

$$\int_{\mathbb{R}^2} |\varphi(t_1,\ t_2)| dt_1 dt_2 < +\infty,$$

alors le couple $\mathbf{X}$ *admet une densité de probabilité* $f_{\mathbf{X}}$ *donnée par*

$$f_{\mathbf{X}}(x_1,\ x_2) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \exp[-i(t_1 x_1 + t_2 x_2)] \varphi(t_1; t_2) dt_1 dt_2.$$

Proposition 22 *Soit* $\mathbf{X} = (X_1,\ X_2)$ *un couple de variables aléatoires de fonction caractéristique* φ.

1. *Si* X_1 *et* X_2 *admettent une espérance, alors les dérivées partielles* $\frac{\partial \varphi}{\partial t_1}$ *et* $\frac{\partial \varphi}{\partial t_2}$ *existent et*

$$\mathbf{E}(X_1) = -i \frac{\partial \varphi}{\partial t_1}(0,\ 0) \text{ et } \mathbf{E}(X_2) = -i \frac{\partial \varphi}{\partial t_2}(0,\ 0).$$

2. *Si* X_1 *et* X_2 *admettent un moment d'ordre 2, alors les dérivées partielles d'ordre 2 de* φ *existent et*

$$\mathbf{E}(X_1 X_2) = -i \frac{\partial^2 \varphi}{\partial t_1 \partial t_2}(0,\ 0); \mathbf{E}(X_1^2) = -i \frac{\partial^2 \varphi}{\partial t_1^2}(0,\ 0);$$

et

$$\mathbf{E}(X_2^2) = -i \frac{\partial^2 \varphi}{\partial t_2^2}(0,\ 0).$$

Proposition 23
Soit $\mathbf{X} = (X_1,\ X_2)$ *un couple de variables aléatoires de fonction caractéristique* φ. *Les variables* X_1 *et* X_2 *sont indépendantes si et seulement si, pour tout* $T \in \mathbb{R}^2$,

$$\varphi(T) = \varphi_{X_1}(t_1)\varphi_{X_2}(t_2).$$

Preuve 9 *On va procéder en deux étapes.*

— *Condition nécessaire*
Supposons que X_1 *et* X_2 *soient indépendantes. Soit* $T \in \mathbb{R}^2$. *Alors on sait que les variables* $e^{it_1X_1}$ *et* $e^{it_2X_2}$ *sont indépendantes. Donc*

$$\begin{aligned}\varphi(T) &= \mathbf{E}(\exp[i(t_1X_1 + t_2X_2)]) \\ &= \mathbf{E}\left(\exp[it_1X_1]\exp[it_2X_2]\right) \\ &= \mathbf{E}\left(\exp[it_1X_1]\right)\mathbf{E}\left(\exp[it_2X_2]\right) \\ &= \varphi_{X_1}(t_1)\varphi_{X_2}(t_2).\end{aligned}$$

— *Condition suffisante*
Le cas réel. Supposons que, pour tout $T \in \mathbb{R}^2$,

$$\varphi(T) = \varphi_{X_1}(t_1)\varphi_{X_2}(t_2).$$

Alors :

$$\begin{aligned}&\int_{\mathbb{R}^2} \exp[i(t_1x_1 + t_2x_2)]\, f_{\mathbf{X}}(x_1,\ x_2)dx_1dx_2 \\ &= \int_{\mathbb{R}} \exp[it_1X_1]f_{X_1}(x_1)dx_1 \int_{\mathbb{R}} \exp[it_2x_2]f_{X_2}(x_2)dx_2 \\ &= \int_{\mathbb{R}^2} \exp[i(t_1x_1 + t_2x_2)]f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2.\end{aligned}$$

Donc $\mathbf{X}$ *a la même fonction caractéristique que le couple de loi* $f_{X_1}f_{X_2}$, *et puisque la fonction caractéristique caractérise la loi, la densité du couple est le produit des densités marginales, et donc* X_1 *et* X_2 *sont indépendantes.*

4.8 Exercices

Exercice 29 *Un poisson pond des œufs au fond du torrent. Leur nombre N suit une loi de Poisson de paramètre $\lambda > 0$. Chaque œuf survit avec une probabilité $p \in]0,1[$, indépendamment des autres.*

1. *Soit M le nombre d'œufs qui survivent. Donner la loi conjointe du couple $(N,\ M)$. Donner la loi marginale et l'espérance de M.*
2. *M et $N - M$ sont-elles indépendantes ?*

Exercice 30 *Soient n et N des entiers supérieurs ou égaux à 2 et $X_1, \ldots, X_n$ des variables aléatoires indépendantes et distribuées uniformément sur l'ensemble $\{1, \ldots,\ N\}$, (i.e. $\mathbf{P}(X_i = k) = \frac{1}{N}$ pour $k = 1, 2, \ldots, N$). On désigne par U_n leur minimum et par V_n leur maximum.*

1. *Calculer la loi de V_n.*
2. *Calculer la loi jointe de U_n et V_n puis $\mathbf{P}(U_n = V_n)$.*

Exercice 31 *Dans le bois de Vincennes, on modélise le diamètre d'un arbre par une variable aléatoire X, et sa hauteur par une autre variable aléatoire Y. La loi jointe de X et Y est donnée par la densité :*

$$f_{X,\ Y}(x, y) = \frac{1}{4}(x + y)e^{-y} \text{ pour } y \geq 0, 0 \leq x \leq 2.$$

1. *Donner la densité marginale de X.*
2. *X et Y sont-elles indépendantes ?*
3. *Calculer $\mathbf{E}[X]$.*
4. *L'âge d'un arbre est donné par $W = 12XY$. Calculer $\mathbf{E}[W]$.*

Exercice 32 *Soit X une variable aléatoire réelle de loi uniforme sur $[0, 1]$.*

1. *Déterminer la loi de $U = \sup\{X, 1 - X\}$.*
2. *Déterminer la loi de $V = \inf\{X, 1 - X\}$.*

Exercice 33 *Soit X et Y deux variables aléatoires indépendantes. Déterminer la loi de la somme $S = X + Y$ si*

1. *X et Y suivent la même loi uniforme sur $[-1, 0]$.*

2. *X suit la loi uniforme sur $[0,1]$ et Y suit la loi uniforme sur $[-3,2]$.*

Exercice 34 *Soit U une variable aléatoire réelle de loi uniforme sur l'intervalle $[0,1]$. On définit la variable aléatoire X par*

$$X = \begin{cases} \sqrt{\frac{1-U}{U}} & si\,;\, U \leq \frac{1}{2} \\ \sqrt{\frac{U}{1-U}} & si\,;\, \frac{1}{2} < x \leq 1 \end{cases}.$$

1. *Déterminer f_X la densité de la loi de X.*
2. *Déterminer F_X la fonction de répartition de la loi de X.*
3. *Déterminer x_0 vérifiant : $F_X(x_0) = 12$ (x_0 s'appelle la médiane de la loi de X).*
4. *Calculer la probabilité suivante : $\mathbf{P}(X \leq 2x_0 | X > x_0)$.*

Exercice 35 *Soit$(X,\ Y)$ un vecteur aléatoire de densité $f(x,y) = \mathbf{1}_{[0,1]^2}(x,y)$. Déterminer les lois de $X,\ Y$ et $Z = XY$.*

Exercice 36 *Soient X et Y deux variables aléatoires indépendantes. X suit une loi $\mathcal{N}(0,1)$ et Y une loi $\mathcal{N}(0,\sigma^2)$, o'u σ désigne un réel positif.*

1. *Écrire la densité de la loi du couple $(X,\ Y)$.*
2. *On pose $U = \frac{Y}{X}$. Calculer la densité de la loi du couple $(X,\ U)$.*
3. *Les variables X et U sont-elles indépendantes ?*

Exercice 37

1. *Soit Z_a une v.a. gamma de paramètre a. Calculer explicitement les moments entiers $\mathbf{E}((Z_a)^n)$, en fonction de a et de $n \in \mathbb{N}$.*
2. *Soient Z_a et Z_b deux variables gamma indépendantes de paramètres respectifs a et b. Montrer que les variables $\frac{Z_a}{(Z_a+Z_b)}$ et $Z_a + Z_b$ sont indépendantes et expliciter la loi de $\frac{Z_a}{(Z_a+Z_b)}$.*

5 Convergence en loi

Le théorème central limite justifie l'importance de la loi normale. En effet, on approxime (en loi) de nombreuses variables aléatoires par des variables aléatoires normales ou des vecteurs gaussiens en dimension $d \geq 2$. Il importe donc de savoir manipuler correctement ce type de convergence. En fait, la définition de la convergence en loi par les fonctions de répartition ne permet de considérer la convergence en loi que pour les variables ou vecteurs aléatoires à valeurs dans un espace fini-dimensionnel (ie. ce qu'on a appelé variables aléatoires réelles ou vecteurs aléatoires).

On entame ce chapitre par deux propositions qui nous serons utiles pour la suite, puis on introduit la convergence en loi pour arriver au théorème central limite.

5.1 Fonction caractéristique et somme de variables

La proposition suivante représente une généralisation de la proposition qui concerne un couple de v. a donnée par la proposition (23).

Proposition 24 *Soit $X_1, \ldots, X_n$ n variables aléatoires indépendantes. Soit $S_n = \sum_{j=1}^{n} X_j$. Alors la fonction caractéristique de S_n est le produit des fonctions caractéristiques des X_j donnée par*

$$\varphi_{S_n}(t) = \prod_{j=1}^{n} \varphi_{X_j}(t), \ \forall t \in \mathbb{R}.$$

Preuve 10

$$\varphi_{S_n}(t) = \varphi_{X_1+\cdots+X_n}(t) = \mathbf{E}\left(\exp^{it\sum_{j=1}^{n} X_j}\right)$$

$$= \mathbf{E}\left(\exp^{\sum_{j=1}^{n} itX_j}\right)$$
$$= \mathbf{E}\left(\prod_{j=1}^{n} \exp^{itX_j}\right).$$

Comme les variables $X_1, \ldots, X_n$ sont indépendantes, d'après une proposition du chapitre précédent, les variables $e^{itX_1}, \ldots, e^{itX_n}$ le sont aussi et donc, d'après la proposition (23),

$$\varphi_{S_n}(t) = \prod_{j=1}^{n} \mathbf{E}\left(\exp^{itX_j}\right) = \prod_{j=1}^{n} \varphi_{X_j}(t).$$

Ce résultat permet d'avoir la proposition suivante.

Proposition 25 *1. Soit $X \simeq \mathcal{B}(n,\ p)$ et $Y \simeq \mathcal{B}(m,\ p)$. Si $X,\ Y$ indépendantes alors*

$$X + Y \simeq \mathcal{B}(n + m,\ p).$$

2. Soit $X \simeq \mathcal{P}(\lambda)$ et $Y \simeq \mathcal{P}(\lambda')$. Si $X,\ Y$ indépendantes alors

$$X + Y \simeq \mathcal{P}(\lambda + \lambda').$$

3. Soit $X \simeq \mathcal{N}(\mu_1; \sigma_X)$ et $Y \simeq \mathcal{N}(\mu_2; \sigma_Y)$. Si $X,\ Y$ indépendantes alors

$$X + Y \simeq \mathcal{N}\left(\mu_1 + \mu_2; \sqrt{\sigma_X^2 + \sigma_Y^2}\right).$$

4. Si X et Y sont deux v.a.r. indépendantes, alors $X+Y$ a pour densité

$$f_{X+Y}(u) = \int_{\mathbb{R}} f_X(u - v) f_Y(v) dv = \int_{\mathbb{R}} f_X(v) f_Y(u - v) dv.$$

Preuve 11 *On démontre seulement les deux derniers points, les autres points sont laissés aux étudiants.*
3. Comme $X \simeq \mathcal{N}(\mu_1; \sigma_X)$, alors $\varphi_X(t) = \exp\left[i\mu_1 t - t^2 \frac{\sigma_X^2}{2}\right]$ et comme $Y \simeq \mathcal{N}(\mu_2; \sigma_Y)$, alors $\varphi_Y(t) = \exp\left[i\mu_2 t - t^2 \frac{\sigma_Y^2}{2}\right]$. Puisque X et Y sont indépendantes, d'après la proposition (24),

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = \exp\left[i(\mu_1 + \mu_2)t - t^2 \frac{\sigma_X^2 + \sigma_Y^2}{2}\right].$$

On reconnaît la fonction caractéristique de la loi normale de paramètres $\mu_1 + \mu_2$ *et* $\sqrt{\sigma_X^2 + \sigma_Y^2}$.
4. Comme X *et* Y *sont indépendantes, d'après la proposition(24),*

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$$

$$= \int_{\mathbb{R}} f_X(x)e^{itx}dx \int_{\mathbb{R}} f_Y(v)e^{itv}dv.$$

5.2 Convergence en loi

Définition 13 *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires définies sur les espaces de probabilités respectifs* $(\Omega_n,\ \mathcal{A}_n,\ \mathbf{P}_n)$ *et* X *une variable aléatoire définie sur* $(\Omega,\ \mathcal{A},\ \mathbf{P})$. *Soit* F_n *la fonction de répartition de* X_n *et* F *celle de* X. *Notons* $C(F)$ *l'ensemble des points où* F *est continue. On dit que la suite* $(X_n)_{n\in\mathbb{N}}$ *converge en loi vers* X *(noté* $X_n \xrightarrow{L} X$*) si, pour tout* $x \in C(F)$,

$$\lim_{n\to+\infty} F_n(x) = F(x).$$

Remarquer que lorsque F_n converge vers une fonction F, il faut vérifier que F est bien une fonction de répartition. Ce n'est pas forcément le cas.

Par contre, la condition suivante permet de définir la convergence en loi pour les variables aléatoires à valeurs dans un espace métrique quelconque composé avec une fonction continue bornée. En général, dans les espaces métriques, on prend donc la relation qui suit pour définition de la convergence en loi.

Théorème 12 *La convergence en loi* $\left[X_n \xrightarrow{L} X\right]$ *est équivalente à avoir pour toute fonction* $g : \mathbb{R} \to \mathbb{R}$ *continue bornée.*

$$\lim_{n\to+\infty} \mathbf{E}[g(X_n)] = \mathbf{E}[g(X)]. \tag{5.1}$$

La démonstration de ce théorème ne sera pas donnée, du fait que l'outil utilisé pour la démonstration n'est disponible qu'on troisième année.

Théorème 13 *Théorème de continuité de Paul-Lévy*
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires de fonction caractéristique φ_n. Soit X une variable aléatoire de fonction caractéristique φ. Alors,

$$\left[X_n \xrightarrow{L} X\right] \Rightarrow \left[\forall t \in \mathbb{R}; \lim_{n\to+\infty} \varphi_n(t) = \varphi(t)\right].$$

Preuve 12 *Le sens direct est immédiat avec le Théorème (12) puisqu'on a $\varphi_X(t) = \mathbf{E}[g_t(X)]$ avec $g_t(x) = e^{itx}$ continue bornée.*
Pour la réciproque, on suppose d'abord que $g : \mathbb{R} \to \mathbb{R}$ est $\mathcal{C}^2$ à support compact. Sa transformée de Fourier $\hat{g}(t) = \int_{\mathbb{R}} e^{itx} g(x)dx$ est alors intégrable et est bornée. En effet, $\hat{g}$ est bornée ainsi que g'' qui est égale à $t^2\hat{g}(t)$. On a donc $\hat{g}(t) = O(1/t^2)$ en $+\infty$ et $\hat{g}$ est bornée. Donc, on peut appliquer le théorème d'inversion de Fourier pour écrire

$$g(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itX_n}\hat{g}(t)dt.$$

Comme $\hat{g}$ est intégrable, le théorème de Fubini (l'intégration successive) s'applique et donne

$$\mathbf{E}[g(X_n)] = \frac{1}{2\pi}\mathbf{E}\left[\int_{\mathbb{R}} e^{-itX_n}\hat{g}(t)dt\right] = \frac{1}{2\pi}\int_{\mathbb{R}} \mathbf{E}\left[e^{-itX_n}\right]\hat{g}(t)dt$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} \varphi_{X_n}(-t)\hat{g}(t)dt.$$

Mais comme $\varphi_{X_n}(-t)$ converge vers $\varphi_X(-t)$ pour chaque t et que $|\varphi_{X_n}(-t)| \leq 1$ pour tout $t \in \mathbb{R}$ avec en plus intégrabilité de $\hat{g}$, par le théorème de convergence dominée, on a

$$\lim_{n\to+\infty} \mathbf{E}[g(X_n)] = \frac{1}{2\pi}\int_{\mathbb{R}} \varphi_X(-t)\hat{g}(t)dt = \mathbf{E}[g(X)].$$

Lorsque l'on suppose seulement g continue à support compact puis ensuite continue bornée, on procède comme précédemment dans la preuve du sens direct de ce Théorème.

Théorème 14 *Théorème centrale limite*
Soit $X_1, X_2, \ldots, X_n$ une suite de v.a indépendantes et de même loi, d'espérance μ et de variance σ^2 et soit la v.a $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Pour n assez

grand, la loi de $Z = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$ *peut être approchée par la loi normale standard* $\mathcal{N}(0,1)$. *C'est-à-dire,*

$$\mathbf{P}\left(\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \leq x\right) \simeq \mathbf{P}\left(Z \leq x\right).$$

Pour un théorème d'une telle importance en statistiques et en probabilité appliquée, il existe une démonstration particulièrement simple utilisant la fonction caractéristique d'une variable aléatoire. Cette démonstration ressemble à celle de la loi des grands nombres.

Preuve 13 *Pour une variable aléatoire* Y *d'espérance 0 et de variance 1, la fonction caractéristique de* Y *admet le développement limité*

$$\varphi_Y(t) = 1 - \frac{t^2}{2} + o(t^2), \quad t \to 0.$$

Si Y *vaut* $\frac{X_i - \mu}{\sigma}$, *il est facile de voir que la moyenne centrée réduite des observations* $X_1, \ldots, X_n$ *est simplement*

$$Z_n = \frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i.$$

D'après les propriétés élémentaires des fonctions caractéristiques, la fonction caractéristique de Z_n *est*

$$\left[\varphi_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right]^n \longrightarrow \mathrm{e}^{-t^2/2}, \text{ lorsque } n \to \infty.$$

Mais cette limite est la fonction caractéristique de la loi normale centrée réduite $\mathcal{N}(0,1)$, *d'où l'on déduit le théorème central limite grâce au théorème de convergence de Lévy (13), qui affirme que la convergence simple des fonctions caractéristiques implique la convergence en loi.*

5.3 Exercices

Exercice 38 *On considère deux suites de v.a. réelles $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$. Donner un exemple dans lequel (X_n) converge en loi vers une v.a. X, (Y_n) converge en loi vers une v.a. Y, mais $(X_n + Y_n)$ ne converge pas en loi. (Indication : $X_n = X$ de loi $\mathcal{N}(0,1)$, et $Y_n = (-1)^n X$).*

Exercice 39 *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles définies sur un espace $(\Omega, \mathcal{A}, \mathbf{P})$, et f une application continue de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que si X_n converge en loi vers X, alors $f(X_n)$ converge en loi vers $f(X)$.*

Exercice 40 *Soit $(Y_n)_{n\geq 1}$ une suite de variables aléatoires. On suppose que $\mathbf{E}[Y_n] \to 1$ et que $\mathbf{E}[(Y_n)^2] \to 1$. Montrer que Y_n converge en loi vers 1.*

Exercice 41 *On considère une suite de v.a. indépendantes et de même loi $(X_n)_{n\geq 0}$. On définit alors la suite $(Y_n)_{n\geq 0}$ par*

$$Y_0 = \frac{X_0}{2},\ Y_1 = \frac{X_1 + Y_0}{2},\ Y2 = \frac{X_2 + Y_1}{2},\ \ldots,\ Y_n = \frac{X_n + Y_{n-1}}{2}.$$

1. *. Calculer la fonction caractéristique φ_n de Y_n en fonction de φ, la fonction caractéristique de X_1, et de n.*
2. *On suppose que la loi commune aux variables X_n est la loi normale centrée $\mathcal{N}(0, \sigma^2)$. Quelle est la loi de Y_n ? Quelle est la loi limite de (Y_n) lorsque n tend vers l'infini ?*
3. *Si les variables X_n suivent la loi de Cauchy de densité $[\pi(1+x^2)]^{-1}$, $x \in \mathbb{R}$, montrer que (Y_n) converge en loi lorsque n tend vers l'infini. Préciser la limite.*

Indication : la fonction caractéristique de la loi de Cauchy est donnée par $\varphi(t) = e^{-|t|}$.

Table de la loi normale centrée réduite ($\Phi(z)$)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Bibliographie

[1] Eva Cantoni, Philippe Huber, and Elvezio Ronchetti. *Maîtriser l'aléatoire. Exercices résolus de probabilités et statistique.* Springer, 2009.

[2] Billingsley, P., (1979). *Probability and measure*. Wiley series in Probability and statistics.

[3] Bochner, S., (1955). *Harmonic Analysis and the Theory of probability.* University of Chicago Press, Chicogo, Illinois.

[4] Gut, A., (2005). *Probability : A graduate course*, Springer tests in statistics.

[5] Jean-Pierre Lecoutre. *Statistique et probabilités.* Dunod, 2002.

[6] Métivier, M., (1972). *Notions fondamentales de la théorie des probabilités, maîtrises de mathématiques*, 2nd edition, Dunod.

[7] Mikael Beatriz-Alkéos Michaïl. *Probabilités et statistiques élémentaires.*

[8] Thérese Phan and Jean-Pierre Rowenczyk. *Exercices et problèmes de Statistique et probabilités*. 2e éd. Dunod, 2012.