Support du cours de Probabilités et Statistiques IUT d'Orléans, Département informatique

Pierre Andreoletti

IUT d'Orléans, Département Informatique
Laboratoire MAPMO (Bât. de Mathématiques UFR Sciences)
email: pierre.andreoletti@univ-orleans.fr

Première partie

Probabilités

Introduction

La théorie des probabilités a pour objectif de modéliser des expériences où plusieurs issues sont possibles, mais où leur réalisation n'est pas déterminée à l'avance (par exemple un lancé de dés), ceci en vue d'évaluer les risques ou de mettre sur pieds des stratégies pour faire face aux aléas. La théorie des probabilités ne va pas permettre de prédire quelle issue va se réaliser, mais quelle chance a chaque issue de se réaliser.

La théorie des probabilités et par extension les statistiques a des applications dans de nombreux domaines :

1. **La biologie et la médecine** : repérage de séquences ADN, tests de médicaments, évaluation du risque de propagation de virus ...
2. **La finance** (Les banques) : évaluation de produits financiers (options ...), maîtrise des risques liés à des investissements ...
3. **Industries** : aéronautique, automobile, chimie ...
4. **Economie** : prévision de croissance, sondages, contrôle de gestion ...
5. **Les jeux** : Loto, Casinos ...

Tous ces champs d'applications sont en lien étroit avec l'informatique : développement et utilisation de logiciels de simulations et de logiciels de traitements statistiques, bases de données ...

Figure 1 – Courbes de fluctuations de deux indices boursiers - Aléatoires ?

Table des matières

Chapitre 1

Evènements Probabilité et Variables aléatoires

Dans un premier temps, on va associer à chaque issue possible un nombre entre 0 et 1 qui traduit notre estimation des chances que cette issue a de se réaliser : on appelle ce nombre la probabilité de cette issue. On appelle “évènement” un ensemble d’issues. La probabilité qu’on associe à un évènement est la somme des probabilités de chaque issue de cet ensemble. Typiquement, la question est de déterminer la probabilité d’un évènement, et la difficulté est d’une part de décrire l’expérience de façon commode afin d’énumérer toutes les issues possibles et leur probabilité respective, et d’autre part de décomposer l’évènement considéré en issues.

1.1 L’espace des issues, l’ensemble des évènements

Avant de calculer les probabilités d’évènements, il faut définir **l’espace des issues** de façon commode et complète. **Cet espace comprendra toutes les issues possibles du jeu, ou de l’expérience aléatoire que l’on considère**, même éventuellement celles qui ne nous intéressent pas, a priori. Dans chaque situation, **l’espace des issues sera noté** Ω (grand omega), alors que **les issues seront notées** ω (petit omega).

Exemple 1.1.1. *On considère un dé à 6 faces, numérotées de 1 à 6. On suppose que le dé est équilibré, ce qui veut dire que les 6 faces ont la même chance de sortir. L’ensemble* Ω *des issues possibles d’un lancer est* $\Omega = \{1, 2, 3, 4, 5, 6\}$. *Une issue possible est* $\{3\}$ *c’est à dire ”la face 3 sort”. Un événement est par exemple ”On obtient un nombre pair” que l’on peut écrire* $\{2, 4, 6\}$.

Exemple 1.1.2. *On considère le même dé, sauf que sur la sixième face, le nombre 6 a été remplacé par 5. Il y a donc deux faces où 5 est inscrit. L’ensemble des issues est ici* $\Omega = \{1, 2, 3, 4, 5\}$. *L’événement ”on obtient un nombre pair” s’écrit* $\{2, 4\}$.

On considère des expériences avec un nombre fini ou dénombrable d’issues. On note Ω cet ensemble d’issues et F **l’ensemble des événements**. F **est l’ensemble des parties de** Ω.

Exemple 1.1.3. *Si* $\Omega = \{1, 2, 3\}$, *alors*

$$F = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}$$

, où $\varnothing$ *est l'ensemble vide.*

Exemple 1.1.4. *Considérons l'exemple 1.1.1, un événement est par exemple :* $\{$*On obtient un nombre pair*$\}$*, que l'on peut aussi écrire de la façon suivante* $\{2, 4, 6\}$*. Pouvez-vous déterminer* F *?*

En théorie des probabilités on a souvent besoin de ramener le calcul de la probabilité d'un évènement au calcul de la probabilité de l'union ou de l'intersection d'événements plus élémentaires, on introduit donc les notations suivantes : Si A et B sont deux sous-ensembles de Ω, on note
$A \cup B = \{\omega; \omega \in A$ **ou** $\omega \in B\}$ **: "A ou B se réalise"**
$A \cap B = \{\omega; \omega \in A$ **et** $\omega \in B\}$ **: "A et B se réalisent"**
$A \setminus B = \{\omega; \omega \in A$ **et** $\omega \notin B\}$ **: "A se réalise mais pas B"**
$\bar{A} = \Omega \setminus A$ **"l'évènement A ne se réalise pas" .**

Deux évènements A et B de F sont disjoints s'ils n'ont aucune issue en commun, c'est à dire que $A \cap B = \varnothing$. Par exemple, A et $\bar{A}$ sont disjoints, ainsi que $\varnothing$ et A.

Exemple 1.1.5. *Prenons l'exemple 1.1.1, soit* $A = \{2, 4, 6\}$*,* $B = \{5, 6\}$*,* $C = \{3\}$*, on a :*
1. $A \cup B = \{2, 4, 5, 6\}$
2. $A \cap B = \{6\}$
3. $A \cap C = \varnothing$
4. $\bar{A} = \{1, 3, 5\}$*,* $\bar{A}$ *correspond à l'évènement "on obtient un nombre impair".*
5. $A \setminus B = \{2, 4\}$ *on remarque que* $A \setminus B = A \cap \bar{B}$*.*
Effectuer une analyse similaire pour l'exemple 1.1.2.

Remarque 1.1.6. Lorsque l'on écrit l'évènement A "on obtient un nombre pair" sous la forme $\{2, 4, 6\}$ il est sous-entendu que les "," sont des "ou", on peut également écrire $A = \{2\} \cup \{4\} \cup \{6\}$.

1.2 Définition et Propriétés d'une probabilité

On commence par définir ce qu'est une probabilité :

Definition 1.2.1. Une probabilité P sur Ω est une fonction de F dans l'intervalle $[0, 1]$ telle que

1. $P(\Omega) = 1$,
2. Si $(A_i)_{i \in I}$ est une famille finie ou dénombrable d'événements deux à deux disjoints (*i.e.* $A_i \cap A_j = \varnothing$ si $i \neq j$), alors

$$P(\cup_{i \in I} A_i) = \sum_{i \in I} P(A_i) \qquad (1.1)$$

De cette définition, découlent les **propriétés fondamentales** suivantes :

Proposition 1.2.2. *1.* $P(\varnothing) = 0$.

2. Pour tout événement A, $P(A) + P(\bar{A}) = 1$.

3. Pour tous événements quelconques A et B,

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

4. Si A et B sont deux événements tels que $A \subseteq B$ alors $P(A) \leq P(B)$.

Exemple 1.2.3. *Reprenons l'exemple 1.1.1, étant donné que le dé est équilibré on pense pour ce cas à une probabilité comme à une fonction p sur Ω, telle que $p(1) + ... + p(6) = 1$ et on associe à chaque issue la probabilité 1/6. En utilisant le second point de la définition d'une probabilité on calcule facilement la probabilité de l'évènement $A =$"on obtient un nombre pair". En effet $A = \{2\} \cup \{4\} \cup \{6\}$, les **3 issues qui composent** A **étant 2 à 2 disjointes**[1], $P(A) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 1/2$. Si on veut maintenant calculer la probabilité de l'évènement "obtenir un nombre impair", on peut utiliser le petit 4. de l'exemple 1.1.5 et la propriété 2. de la Proposition 1.2.2*

Pour l'exemple 1.1.2, comme le dé est équilibré, les faces 1, 2, 3 et 4 ont chacune la probabilité 1/6 de sortir, alors que $P(5) = 2/6$. Quelle est la probabilité qu'un nombre pair sorte ?

L'exemple suivant a pour but de manipuler la propriété 3. de la Proposition 1.2.2

Exemple 1.2.4. *Toujours sur l'exemple 1.1.1, soit $A = \{2, 4, 6\}$, $B = \{5, 6\}$, $C = \{3\}$, calculons de deux façon différentes :*
1. On se rappelle du 1. de l'exemple 1.1.5 et on a donc $P(A \cup B) = P(\{2, 4, 5, 6\})$ et comme chacun des évènements $\{2\}$, $\{4\}$, $\{5\}$ et $\{6\}$ sont deux à deux disjoints $P(A \cup B) = P(2) + P(4) + P(5) + P(6) = 4/6 = 2/3$.
2. On utilise la propriété 3. de la Proposition 1.2.2. On a vu dans l'exemple précédent que $P(A) = 1/2$, il est facile de voir que $P(B) = 1/6 + 1/6 = 1/3$ et que d'aprés le petit 2. de l'Exemple 1.1.5, $A \cap B = \{6\}$ et donc que $P(A \cap B) = 1/6$. Ainsi on obtient $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1/2 + 1/3 - 1/6 = 2/3$.

1.3 Variables Aléatoires Discrètes

1.3.1 Définition et Notations

Definition 1.3.1. Une fonction sur Ω à valeurs réelles est appelée variable aléatoire.

Notation Les variables aléatoires seront notées par des lettres majuscules, $X, Y, ...$. Les valeurs qu'elles prennent lorsqu'une issue ω se réalise sont notées par des lettres minuscules. Par exemple on pourra écrire x à la place de $X(\omega)$.

1. *en effet on ne peut obtenir qu'un unique chiffre lors d'un lancé !*

Exemple 1.3.2. *Donnons un premier exemple élémentaire de variable aléatoire. Une variable de **Bernoulli** de paramètre* $p \in [0,1]$ *est une variable aléatoire qui prend soit la valeur* 0 *avec une probabilité* $1-p$ *soit la valeur* 1 *avec une probabilité* p*, on l'utilise souvent pour décrire une expérience aléatoire ayant 2 issues possibles. Un autre exemple, que l'on reprendra en détail en TD, on considère le jeté de deux dés simultanément. On peut définir une variable aléatoire qui est égale à la somme des deux dés.*

1.3.2 Loi d'une variable aléatoire

Definition 1.3.3. Soit P une probabilité sur un espace des issues Ω. Soit X une variable aléatoire définie sur Ω. Lorsqu'à chaque valeur $x_i (1 \leq i \leq n)$ de X on associe les probabilités p_i de l'événement "$X = x_i$", on dit que l'on définit la loi de probabilité P_X de la variable aléatoire X.

Remarque 1.3.4. Pour connaître la loi d'une variable aléatoire, il faut connaître l'ensemble de ses valeurs possibles et la probabilité avec laquelle elle réalise chaque valeur.

Exemple 1.3.5. *Une variable de **Bernoulli*** X *de paramètre* p *a pour loi :*

Valeur de $X \equiv x_i$	*0*	*1*
$p_i \equiv P(X = x_i)$	*1-p*	*p*

où p *est compris entre 0 et 1. Ce tableau se lit de la façon suivante :* $p_0 \equiv P(X = 0) = 1 - p$ *et* $p_1 \equiv P(X = 1) = p$*. Par convention on notera* $X \sim B(p)$ *lorsque* X *suit une loi de Bernoulli. On remarque que si* $p = 1$ *alors* X *est constante égale à* 0*.*

Prenons maintenant un exemple un peu plus compliqué,

Exemple 1.3.6. *on considère 2 jetés successifs d'une pièce de monnaie équilibrée. On définit une variable aléatoire* X *qui compte le nombre de face que l'on a obtenu. Sur deux lancés de pièces* X *peut prendre **soit** la valeur* 0 *qui correspond à l'évènement "la pièce n'est jamais tombé sur face", **soit** la valeur* 1 *qui correspond au cas où "la pièce est tombé une fois sur face" soit la valeur 2 si on obtient deux fois face. Etant donné que la pièce est équilibrée on obtient pour* X *la loi suivante :*

x_i	*0*	*1*	*2*
$P(X = x_i)$	*1/4*	*1/2*	*1/4*

On remarque bien entendu que $P(X = 0) + P(X = 1) + P(X = 2) \equiv p_0 + p_1 + p_2 = 1$*.*

Remarque 1.3.7. Bien que très élementaire ce dernier exemple sous entend l'indépendance des deux lancés. La notion d'**indépendance** qui est intuitive ici sera explicité dans le chapitre suivant. On pourrait par exemple supposer que l'on lance une seconde fois la pièce que si on a obtenu face au premier coup, les deux lancés ne sont alors plus indépendants et la loi de X est donnée par :

x_i	0	1	2
$P(X = x_i)$	1/2	1/4	1/4

1.3.3 Fonction de répartition

Proposition 1.3.8. *La loi d'une variable aléatoire X est caractérisée par sa fonction de répartition F_X définie par :*

$$F_X: \quad \mathbb{R} \to [0,1]$$
$$x \mapsto P(X \leq x)$$

Remarque 1.3.9. F_X est par définition une fonction en escalier, qui est constante entre deux valeurs successives prises par la variable aléatoire X et qui fait un saut en chacune de ses valeurs. On retrouve donc sur le graphe de F_X d'une part les valeurs prises par X ainsi que les probabilités d'obtenir les valeurs.
Il est équivalent de connaître la loi de X ou sa fonction de répartition.

Exemple 1.3.10. *Reprenons l'exemple 1.3.6, la fonction de répartition de X est donnée par le graphe suivant :*

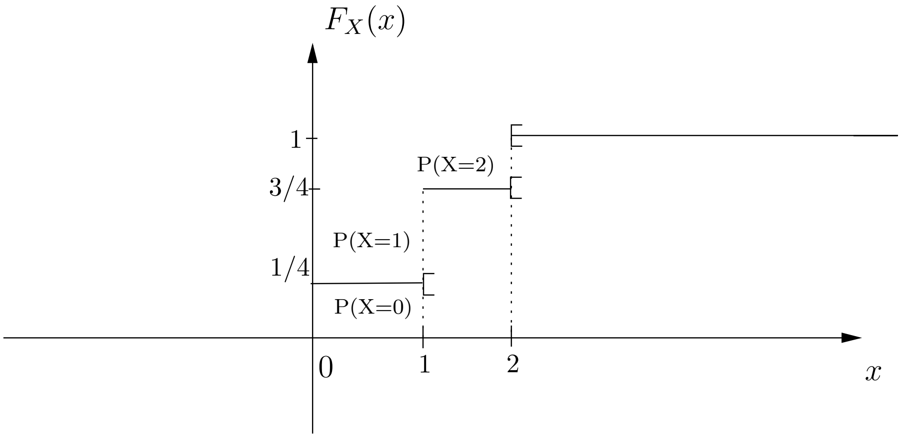


FIGURE 1.1 – Fonction de répartition pour l'Exemple 1.4.6

1.3.4 Fonction indicatrice

On définit dans ce paragraphe la fonction indicatrice, et on donne quelques propriétés.

Definition 1.3.11. La fonction indicatrice notée $\mathbb{1}_A$ d'un événement $A \subset \Omega$, est la variable aléatoire donnée par

$$\mathbb{1}_A: \quad \Omega \to \{0,1\}$$
$$w \mapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

Proposition 1.3.12. *Pour tout $A \subset \Omega$ et $B \subset \Omega$, on a :*

1. $\mathbb{1}_{A\cap B} = \mathbb{1}_A * \mathbb{1}_B$,
2. *Si* $A \cap B = \varnothing$ *alors* $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B$,
3. $\mathbb{1}_{\bar{A}} = 1 - \mathbb{1}_A$.

Remarque 1.3.13. La fonction indicatrice $\mathbb{1}_A$ est une variable aléatoire de Bernoulli de paramètre $p = P(A)$.

Chapitre 2

Indépendance et probabilités conditionnelles

2.1 Indépendance

2.1.1 Cas de deux évènements ou deux variables aléatoires

On commence par donner la définition de deux événements indépendants :

Definition 2.1.1. Deux événements A et B sont indépendants si et seulement si

$$P(A \cap B) = P(A)P(B). \tag{2.1}$$

Exemple 2.1.2. *On s'intéresse à deux lancés d'un dé. On suppose que ces deux lancés sont indépendants, c'est à dire que le résultat du lancé du premier dé n'a aucune influence sur le lancé du second dé. Ainsi les événements $A = \{On\ obtient\ 2\ au\ premier\ lancé\}$ et par exemple $B = \{On\ obtient\ 5\ au\ second\ lancés\}$ sont indépendants. En terme de probabilité cela se traduit par :*

$$P(A \cap B) = P(A)P(B) = (1 \setminus 6) * (1 \setminus 6). \tag{2.2}$$

On définit maintenant l'indépendance pour deux variables aléatoires :

Definition 2.1.3. Deux variables aléatoires $X : \Omega \to \{x_1, \cdots, x_M, \cdots\}$ et $Y : \Omega \to \{y_1, \cdots, x_N, \cdots\}$ sont indépendantes si pour tous i et j,

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j). \tag{2.3}$$

Exemple 2.1.4. *On s'intéresse à deux variables aléatoires de Bernoulli : X_1 de paramètre p et X_2 de paramètre q. On suppose que X_1 et X_2 sont indépendantes. Le couple (X_1, X_2) peut prendre les valeurs $(0, 0)$; $(0, 1)$, $(1, 0)$ et $(1, 1)$. Par indépendance on a*

$$P((0, 0)) \equiv P(X_1 = 0, X_2 = 0) = P(X_1 = 0)P(X_2 = 0) = (1 - p) * (1 - q). \tag{2.4}$$

Maintenant définissons la variable aléatoire $Y = X_1 + X_2$, on peut calculer la probabilité $P(Y = 0)$ de la façon suivante : On remarque que l'événement $\{Y = 0\} = \{X_1 = 0, X_2 = 0\}$ on en déduit donc, d'après ce qui précède que $P(Y = 0) = (1-p) * (1-q)$. Pouvez vous calculer $P(Y = 1)$ et $P(Y = 2)$?

Remarque 2.1.5. Dans l'exemple précédent si on suppose de plus que $q = p$, la variable Y est connue sous le nom de **variable Binomiale** de **paramètres** p **et** 2 correspondant au fait que l'on a sommé 2 variables de Bernoulli de même paramètre p indépendantes pour obtenir Y. Nous reviendrons sur les variables Binomiales dans le chapitre suivant.

2.1.2 Cas d'un nombre fini d'évènements ou variables aléatoires

On généralise la notion d'indépendance à un nombre fini d'évènements et de variables aléatoires :

Definition 2.1.6. Les événements $\{A_1, A_2, \cdots, A_n\}$ sont indépendants si et seulement si pour tout ensemble d'indices $I \subset \{1, \cdots, n\}$

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i), \tag{2.5}$$

on rappelle que si $I = \{i_1, \cdots, i_k\}$, $\prod_{j \in I} P(A_j) \equiv P(A_{i_1}) * P(A_{i_2}) * \cdots * P(A_{i_k})$.

Exemple 2.1.7. *Si on a un ensemble de 3 évènements A_1, A_2 et A_3, ils sont indépendants si et seulement si :*

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3), \tag{2.6}$$
$$P(A_1 \cap A_2) = P(A_1)P(A_2), \tag{2.7}$$
$$P(A_1 \cap A_3) = P(A_1)P(A_3), \tag{2.8}$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3). \tag{2.9}$$

Si on lance 3 fois successivement et indépendamment une pièce de monnaie et que l'on souhaite calculer la probabilité de l'événement $\{P, F, P\} \equiv \{$on obtient Pile au premier lancé,Face au second et Pile au dernier$\}$, par indépendance on a :

$$P(\{P, F, P\}) = P(P)P(F)P(P) = 1/8. \tag{2.10}$$

Quelle est la probabilité que l'on obtienne 2 fois Pile ?

Indépendance pour une suite de variables aléatoires :

Definition 2.1.8. Les variables aléatoires $\{X_1, X_2, \cdots, X_n\}$ sont indépendantes si pour tout ensemble d'indices $I \subset \{1, \cdots, n\}$ et tous réels x_i appartenant aux valeurs possibles de X_i, $(i \in I)$:

$$P\left(\bigcap_{i \in I} \{X_i = x_i\}\right) = \prod_{i \in I} P(X_i = x_i). \tag{2.11}$$

Exemple 2.1.9. *Supposons que pour tout $1 \leq i \leq 8$, X_i suit une loi de Bernoulli de paramètre p et supposons que la suite $\{X_1, X_2, \cdots, X_8\}$ soit indépendante, soit $I = \{1, 3, 6, 8\}$ on a*

$$\begin{aligned} P\left(\bigcap_{i\in I}\{X_i = 1\}\right) &\equiv P(X_1 = 1, X_3 = 1, X_6 = 1, X_8 = 1) \\ &= P(X_1 = 1) * P(X_3 = 1) * P(X_6 = 1) * P(X_8 = 1) = p^4. \quad (2.12) \end{aligned}$$

Soit $I = \{1, 3\}$ et $J = \{5, 6\}$ on a

$$\begin{aligned} P\left(\bigcap_{i\in I}\{X_i = 1\}\bigcap_{j\in I}\{X_j = 0\}\right) &\equiv P(X_1 = 1, X_3 = 1, X_5 = 0, X_6 = 0) \\ &= P(X_1 = 1) * P(X_3 = 1) * P(X_5 = 0) * P(X_6 = 0) \\ &= p^2(1-p)^2. \quad (2.13) \end{aligned}$$

On termine ce paragraphe par le résultat suivant :

Proposition 2.1.10. *Soient X et Y deux variables aléatoires indépendantes. Si f et $g : \mathbb{R} \to \mathbb{R}$ sont deux fonctions quelconques, alors $f(X)$ et $f(Y)$ sont indépendantes.*

2.2 Probabilités conditionnelles

2.2.1 Définitions et propriétés

On commence par une définition

Definition 2.2.1. Soient A et B deux événements avec $P(B) > 0$. La probabilité conditionnelle de A sachant B est donnée par

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

On dit que $P(A|B)$ est la probabilité de A sachant B.

Exemple 2.2.2. *On reprend ici l'exemple de la Remarque 1.3.7. On effectue 2 lancés d'une pièce de monnaie mais avec la règle suivante : on ne relance la pièce que si l'on a obtenu face au premier lancé. Il est intuitif que si l'on obtient face au premier lancé alors on obtient face au second avec une probabilité 1/2. Notons $A = \{$On obtient Face au second lancé$\}$ et $B = \{$On obtient face au premier lancé$\}$. On remarque que $P(A) = P(A\cap B)+P(A\cap \bar{B}) = P(A\cap B) =$1/4 de plus $P(B) = 1/2$ d'où*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 1/2. \quad (2.14)$$

On donne maintenant le résultat suivant que l'on vérifie ensuite sur un exemple simple :

Proposition 2.2.3. *Soient A et B deux événements indépendants et tel que $P(B) > 0$, alors*

$$P(A|B) = P(A) \tag{2.15}$$

Exemple 2.2.4. *On utilise les mêmes notations que pour l'exemple précédent. Si on s'intéresse à deux lancés indépendants d'une pièce de monnaie, indépendant signifiant que quel que soit le résultat du premier lancé on lance une seconde fois la pièce, on a donc $P(A) = P(B) = 1/2$ de plus $P(A \cap B) = 1/4$, ainsi*

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = 1/2. \tag{2.16}$$

On a donc bien $P(B|A) = P(B)$.

2.2.2 Probabilité totale et Théorème de Bayes

Comme nous l'avons déjà mentionné au début de ce cours il est parfois utile de décomposer un événement en sous ensemble d'évènements élémentaires afin d'en calculer la probabilité avec plus de facilité. La proposition suivante va dans ce sens.
On commence par donner la définition d'une partition de Ω

Definition 2.2.5. On appellera *partition de* Ω toute suite $(A_i, i \in I)$ vérifiant $A_i \cap A_j = \phi$ pour tout $i \neq j$ et $\sum_{i \in I} P(A_i) = 1$.

Proposition 2.2.6. ***Formule des probabilités totales*** . *Soit $(A_i, i \in I)$ une partition finie ou dénombrable de Ω, telle que pour tout i, $P(A_i) > 0$. Pour tout événement $B \subset \Omega$,*

$$P(B) = \sum_{i \in I} P(B|A_i)P(A_i). \tag{2.17}$$

Exemple 2.2.7. *Supposons qu'on dispose de deux urnes contenant des boules blanches et noires. La première urne contient 100 boules, dont 99 blanches. La deuxième urne contient 100 boules, dont 10 blanches. Un jeu consiste à choisir une urne au hasard et à tirer une boule dans cette urne. Notons B l'événement "On a tiré une boule blanche", on va donner une partition de cet évènement. Soient $A_1 = \{On\ a\ tir\acute{e}\ dans\ la\ premi\grave{e}re\ urne\}$ et $A_2 = \{On\ a\ tir\acute{e}\ dans\ la\ deuxi\grave{e}me\ urne\}$, on a $\Omega = A_1 \cup A_2$ de plus $A_1 \cap A_2 = \phi$, on peut donc écrire que $B = B \cap A_1 \cup B \cap A_2$, et ainsi*

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) \\ &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2). \end{aligned}$$

Il est facile de voir que $P(A_1) = 1/2$ et $P(A_2) = 1/2$, de plus $P(B|A_1) = 99/100$ et $P(B|A_2) = 1/10$. On en déduit donc que $P(B) = 109/200$.

Proposition 2.2.8. ***Formule de Bayes***. *Soit $(A_i, i \in I)$ une partition finie ou dénombrable de Ω , et soit B tel que $P(B) > 0$. Pour tout i, on a*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j \in I} P(B|A_j)P(A_j)}. \tag{2.18}$$

Exemple 2.2.9. *On reprend l'exercice précédent, et l'on cherche à calculer* $P(A_1|B)$*, par la formule de Bayes on a :*

$$P(A_1|B) \quad = \quad \frac{P(A_1) * P(B|A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)}$$

D'après l'exemple précédent $P(B) = 109/200$ *et* $P(B|A_1) = 99/100$*, on en déduit* $P(A_1|B) = 99/109$*.*

Chapitre 3

Variables Aléatoires Discrètes - Moyenne - Variance

Dans ce chapitre on va introduire de nouvelles notions pour les variables aléatoires. Ces nouvelles notions s'appuient sur les définitions et exemples du paragraphe 1.3.

3.1 Espérance (Moyenne) d'une variable aléatoire discrète

La moyenne ou espérance est une valeur qui sert à avoir une idée de la valeur typique d'une variable aléatoire. Elle n'est en générale pas suffisante pour comprendre comment se comporte cette variable aléatoire mais donne une première indication. Son avantage est qu'elle est très facile à calculer dans la majorité des cas.

3.1.1 Définition et premiers exemples

Commençons par donner une définition :

Definition 3.1.1. Soit Ω un espace fini ou dénombrable, P une probabilité sur Ω, et X une variable aléatoire. On appelle **espérance** de X (ou **moyenne** de X) la quantité

$$E(X) = \sum_{\omega \in \Omega} X(\omega) P(\omega). \tag{3.1}$$

Remarque 3.1.2. X admet une moyenne que si $\sum_{\omega \in \Omega} X(\omega) P(\omega)$ existe.

Commençons par un premier exemple général très utile dans la pratique,

Exemple 3.1.3. *Considérons la variable X ayant pour loi :*

Valeurs X	x_1	x_2	$\cdots$	x_n
$P(X = x_i)$	p_1	p_2	$\cdots$	p_n

alors $E(X) = \sum_{i=1}^{n} p_i x_i \equiv p_1 x_1 + p_2 x_2 + \cdots + p_n x_n$.

Voyons maintenant un exemple pratique

Exemple 3.1.4. *Considérons le jeu suivant, on lance un dé et on définit la variable aléatoire X qui prend la valeur de la face du dé, la loi de X est*

Valeurs de X	1	2	3	4	5	6
Probabilités $P(X = x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

par l'exemple précédent on en déduit $E(X) = 1*1/6+2*1/6+3*1/6+4*1/6+5*1/6+6+1/6 = 3$.

Remarque 3.1.5. Dans l'exemple précédent on remarque que la **probabilité d'obtenir une valeur (entre 1 et 6) est toujours la même** égale à 1/6. **On dit que X suit une loi uniforme**.

Deux exemples fondamentaux :

Exemple 3.1.6. ***Moyenne d'une variable de Bernoulli****. Soit X une variable de Bernoulli de paramètre p, par définition d'une variable de Bernoulli (Exemple 1.4.5), on a*

$$E(X) \;=\; 1 * P(X = 1) + 0 * P(X = 0) = p.$$

Ainsi l'espérance d'une variable de Bernoulli de paramètre p ***est égale à*** p.

Exemple 3.1.7. ***Moyenne d'une fonction indicatrice****. Soit A un sous ensemble de Ω, on cherche à calculer $E(\mathbb{1}_A)$, on a d'après la définition :*

$$\begin{aligned} E(\mathbb{1}_A) &= \sum_{\omega\in\Omega} \mathbb{1}(\omega)P(\omega) \\ &= \sum_{\omega\in A} \mathbb{1}(\omega)P(\omega) + \sum_{\omega\in\bar{A}} \mathbb{1}(\omega)P(\omega) \\ &= 1 * P(A) + 0 * P(\bar{A}) = P(A). \end{aligned}$$

Ainsi l'espérance de $\mathbb{1}_A$ ***est égale à*** $P(A)$.

3.1.2 Propriétés de l'espérance (moyenne)

On énonce maintenant les propriétés élémentaires de l'espérance, dans toute la suite X et Y sont deux variables aléatoires.

Proposition 3.1.8. *Linéarité de l'espérance. Soit $a \in \mathbb{R}$, alors*

1. *E(aX)=aE(X),*
2. *E(X+Y)=E(X)+E(Y).*

D'autres propriétés élémentaires que l'on utilise fréquemment sont

Proposition 3.1.9. *Soit $a \in \mathbb{R}$, alors*

1. *E(a)=a,*
2. $X \geq 0 \Rightarrow E(X) \geq 0$,
3. $X \geq Y \Rightarrow E(X) \geq E(Y)$.

Exemple 3.1.10. *Soient X et Y deux variables de Bernoulli de paramètre respectif p et q. Pour calculer $E(X+Y)$ il suffit de sommer $E(X)$ et $E(Y)$, dans l'exemple 3.1.5 on a vu que la moyenne d'une variable de Bernoulli est égale à son paramètre, on en déduit $E(X+Y) = p+q$. Pour tout $a \in \mathbb{R}$, on a également $E(aX) = ap$.*

Pour terminer on énonce une **propriété de l'espérance** qui est **vraie** que si l'on a **indépendance** entre variables aléatoires :

Proposition 3.1.11. ***Si** X **et** Y **sont indépendantes**, alors :*

$$E(XY) = E(X)E(Y). \tag{3.2}$$

Exemple 3.1.12. *Prenons deux variables de Bernoulli indépendantes X et Y de paramètre respectif p et q, alors $E(XY) = p * q$. Supposons maintenant que $Y = \mathbb{1}_A$ ou $A \subset \Omega$, alors on a $E(XY) = E(X) * E(Y) = p * P(A)$.*

3.2 La Variance et la Covariance

Comme nous l'avons dit précédemment la moyenne n'est pas suffisante pour avoir une bonne idée du comportement d'une variable aléatoire. La variance mesure un écart (quadratique) moyen de la variable aléatoire par rapport à cette moyenne. On dit qu'elle mesure une dispersion de la variable par rapport à sa moyenne.
On commence par définir le **second moment** d'une variable aléatoire

Definition 3.2.1. Soit Ω un espace fini ou dénombrable, P une probabilité sur Ω, et X une variable aléatoire. On appelle **second moment** de X la quantité

$$E(X^2) = \sum_{\omega \in \Omega} X^2(\omega)P(\omega). \tag{3.3}$$

Remarque 3.2.2. $E(X^2)$ n'existe que si $\sum_{\omega \in \Omega} X^2(\omega)P(\omega)$ existe.

Definition 3.2.3. La variance d'une variable aléatoire X, telle que $E(X^2) < \infty$, et notée $var(X)$ est le nombre positif suivant

$$var(X) = E((X - E(X))^2) \tag{3.4}$$

Remarque 3.2.4. Dans la pratique on utilise la "définition" suivante qui est en fait une simple conséquence de la définition précédente :

$$var(X) = E(X^2) - (E(X))^2. \tag{3.5}$$

Donnons un premier exemple général

Exemple 3.2.5. *Reprenons la variable X de l'Exemple 3.1.2, par définition on a :*

$$E(X^2) = \sum_{i=1}^{n} x_i^2 p_i, \tag{3.6}$$

$$Var(X) = \sum_{i=1}^{n} x_i^2 p_i - (\sum_{i=1}^{n} x_i p_i)^2. \tag{3.7}$$

Calculons la variance d'une variable de Bernoulli de paramètre p, on a : $E(X^2) = 1^2 * p + 0^2 * (1-p) = p$ ainsi $var(X) = p - p^2$.

Donnons immédiatement quelques propriétés de la variance

Proposition 3.2.6. *Soit X une variable aléatoire et $a \in \mathbb{R}$,*

1. *$var(aX) = a^2 var(X)$,*
2. *$var(X + a) = var(X)$,*
3. *$var(a) = 0$,*
4. *$var(X) = 0$ ssi pour tout $\omega \in \Omega$ tel que $P(\omega) > 0$, $X(w) = E(X)$.*

Donnons maintenant une propriété de la variance pour deux variables aléatoires indépendantes.

Proposition 3.2.7. *Soient X **et** Y **deux variables aléatoires indépendantes**, on a*

$$var(X + Y) = var(X) + var(Y). \tag{3.8}$$

On introduit maintenant une nouvelle valeur la **covariance** elle mesure la corrélation entre deux variables aléatoires.

Definition 3.2.8. Soient X et Y deux variables aléatoires, on a

$$Cov(X, Y) = E((X - E(X))(Y - E(Y)). \tag{3.9}$$

Proposition 3.2.9. *Soient X et Y deux variables aléatoires, on a*

1. *Cov(X,Y)=E(XY)-E(X)E(Y),*
2. ***Si** X et Y sont indépendantes **alors** $Cov(X, Y) = 0$.*

Exemple 3.2.10. *Prenons $X \sim B(p)$, et $Y = -X$, calculons $Cov(X, Y)$ on a :*

$$E(XY) = -E(X^2) = -p, \tag{3.10}$$

on sait de plus que $E(X)E(Y) = -(E(X))^2 = -p^2$, d'après la Proposition ci-dessus on en déduit $Cov(X, Y) = -p(1 - p)$.
Supposons maintenant que X et Y sont 2 variables aléatoires quelconques indépendantes, d'après la Proposition 3.1.10 $E(XY) = E(X)E(Y)$ donc $Cov(X, Y) = 0$.

Remarque 3.2.11. Il est facile de montrer que pour tout X et Y, $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$.

3.3 Lois classiques

3.3.1 Loi de Bernoulli $B(p)$

Nous avons déjà beaucoup parlé des variables de Bernoulli, c'est une variable aléatoire très élémentaire prenant deux valeurs différentes. Dans ce paragraphe nous résumons ses principales propriétés.
Soit $X \sim B(p)$, la loi de X est donnée par $P(X = 1) = p$ et $P(X = 0) = 1 - p$.

Proposition 3.3.1.

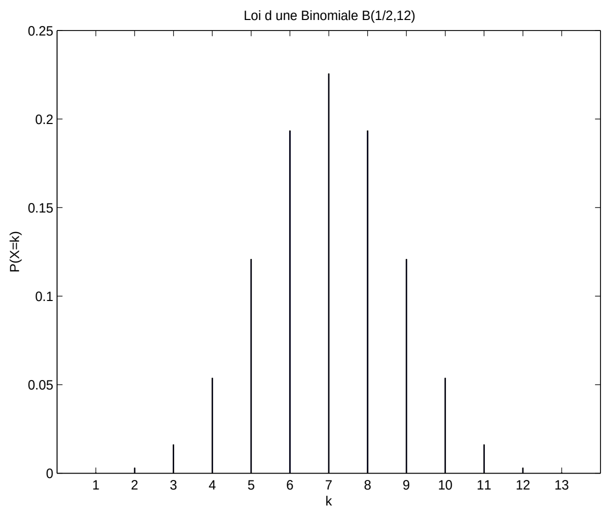
$$E(X) = p,$$
$$Var(X) = p(1 - p).$$

3.3.2 Loi Binomiale $B(n, p)$

Definition 3.3.2. Soient $(X_1, X_2, \cdots, X_n)$ $n \in \mathbb{N}$ variables aléatoires indépendantes de Bernoulli, chacune de ces variables ayant pour paramètre p, la variable aléatoire S_n définie par $S_n = \sum_{i=1}^{n} X_i$ suit une loi binomiale de paramètres n et p, **on note** $S_n \sim B(n, p)$. De plus la loi de S_n est donnée par : pour tout $k \in \{1, 2, \cdots, n\}$

$$P(S_n = k) = C_n^k p^k (1 - p)^{n-k} \tag{3.11}$$

où $C_n^k = \frac{n!}{(n-k)!k!}$. On rappelle que $k!$ (factoriel k) est égale à $k * (k - 1) * (k - 2) * \cdots * 2 * 1$.



On donne maintenant la moyenne et la variance d'une Binomiale de paramètres p et n

Proposition 3.3.3.

$$E(X) = n * p, \tag{3.12}$$
$$Var(X) = np(1 - p). \tag{3.13}$$

Un exemple typique où une variable binomiale apparaît est le suivant :

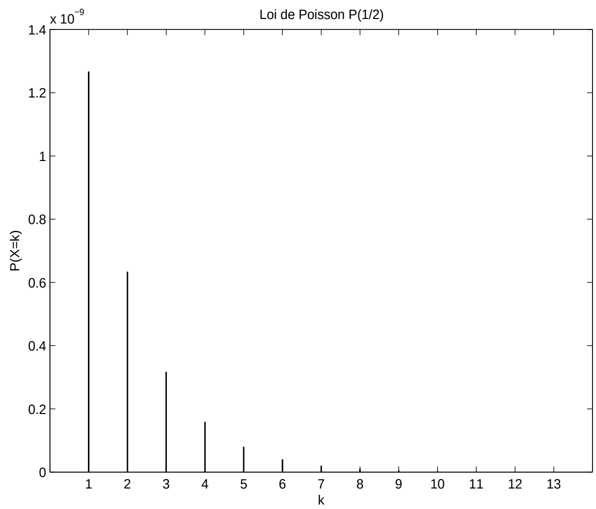
Exemple 3.3.4. *On effectue n lancés d'une pièce de monnaie équilibrée, on définit S_n la variable aléatoire qui va compter le nombre de fois ou la pièce tombe sur $Face$. Pour comprendre pourquoi S_n suit une loi binomiale on effectue la démarche suivante : on définit X_1 la variable aléatoire qui prend la valeur 1 si la pièce tombe sur "face" lors du premier lancé et 0 sinon, de même on définit pour tout $2 \leq i \leq n$ la variable aléatoire X_i qui prend la valeur 1 si la pièce tombe sur "face" lors du $i^{ème}$ lancé. Il apparaît clairement que chaque variable X_i suit une loi de Bernoulli de paramètre 1/2. $\sum_{i=1}^{n} X_i$ compte le nombre de face que l'on a obtenu sur n lancés. Chaque lancé étant indépendant les uns des autres $S_n = \sum_{i=1}^{n} X_i$ suit une loi binomiale de paramètre 1/2 et n.*

3.3.3 Loi de Poisson $P(\lambda)$

Commençons par une définition qui donne la loi d'une variable de Poisson :

Definition 3.3.5. X suit une loi de Poisson de paramètre λ si et seulement si pour tout $k \in \mathbb{N}$ on a :

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}. \tag{3.14}$$



Remarque 3.3.6. Cette variable aléatoire est différente de toutes celles qui ont été introduites jusqu'à maintenant car elle prend un nombre infini dénombrable de valeurs. On peut vérifier que 3.14 est une mesure de probabilité en montrant que :

$$\sum_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} = 1. \tag{3.15}$$

La preuve est élémentaire, elle découle de la définition de e^{λ}.

On donne maintenant la moyenne et la variance de X,

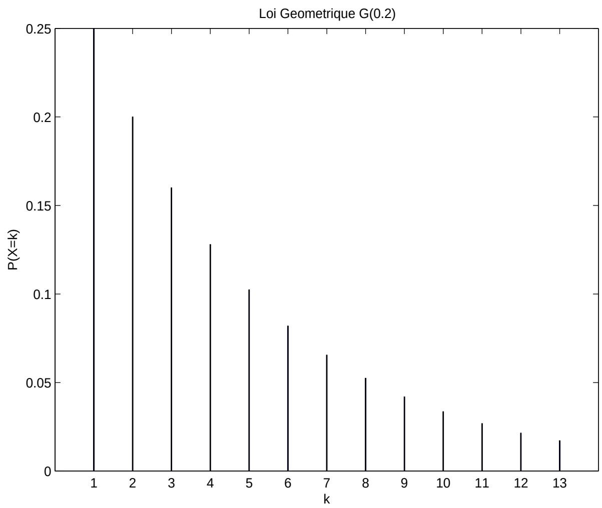
Proposition 3.3.7. *Si* $X \sim P(\lambda)$

$$E(X) = \lambda, \tag{3.16}$$
$$Var(X) = \lambda. \tag{3.17}$$

3.3.4 Loi Géométrique $G(p)$

Definition 3.3.8. X suit une loi de géométrique de paramètre p si et seulement si pour tout $k \in \mathbb{N}$ on a :

$$P(X = k) = (1-p)^{k-1}p. \tag{3.18}$$



De même que la loi de Poisson une variable aléatoire qui suit une loi Géométrique prend un nombre infini dénombrable de valeur avec une probabilité strictement positive. On vérifie que $\sum_{k=0}^{+\infty}(1-p)^{k-1}p = 1$.

Proposition 3.3.9. *Si* $X \sim G(p)$

$$E(X) = \frac{1}{p}, \tag{3.19}$$
$$Var(X) = \frac{1-p}{p^2}. \tag{3.20}$$

3.4 Fonction génératrice

La notion de fonction génératrice, comme la fonction de répartition vue au paragraphe 1.4.3 est une façon de caractériser la loi d'un variable aléatoire. C'est à dire qu'il est équivalent de connaître une loi de probabilité ou sa fonction génératrice.

Definition 3.4.1. Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La fonction génératrice G_X de X est donnée par

$$G_X(t) = E(t^X) = \sum_{k=0}^{+\infty} t^k P(X = k). \tag{3.21}$$

Remarque 3.4.2. De façon générale, G est définie pour tout $t \in [-1, +1]$, $G_X(t)$ existe si et seulement si la série de terme générale $t^k P(X = k)$ est absolument convergente.

Proposition 3.4.3. *La fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$ caractérise sa loi. Autrement dit, si X et Y sont deux variables à valeurs dans $\mathbb{N}$,*

$$G_X = G_Y \iff \forall k \in \mathbb{N}, P(X = k) = P(Y = k). \tag{3.22}$$

Exemple 3.4.4. *Soit X une variable de Bernoulli de paramètre p, on a :*

$$G_X(t) = t^0 P(X = 0) + t^1 P(X = 1) = (1 - p) + pt. \tag{3.23}$$

On énonce maintenant un résultat pour une somme de variables indépendantes :

Proposition 3.4.5. *Soient X et Y des variables aléatoires indépendantes, alors*

$$G_{X+Y}(t) = G_X(t) * G_Y(t). \tag{3.24}$$

Plus généralement,

Proposition 3.4.6. *Soient $X_1, X_2 \cdots X_n$ des variables aléatoires indépendantes, alors*

$$G_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^{n} G_{X_i}(t). \tag{3.25}$$

La fonction génératrice permet de retrouver facilement la moyenne et le second moment d'une variable aléatoire :

Proposition 3.4.7. *Soit X une variable aléatoire, alors*

$$G'_X(1) = E(X), G''_X(1) = E(X^2) - E(X), \tag{3.26}$$

où G'_X désigne la dérivée première de G_X et G''_X sa dérivée seconde.

Chapitre 4

Variables Aléatoires Continues

Les variables aléatoires que nous avons étudiées jusqu'à maintenant prenaient un nombre fini ou dénombrable de valeurs. Il est très souvent intéressant, en particulier dans les applications industrielles, de considérer des variables aléatoires à valeurs continues dans un espace infini non dénombrable. Par exemple, en finance on modélise le prix d'une action par une variable aléatoire à valeurs continues dans $\mathbb{R}^+$. En physique on peu modéliser la trajectoire d'une particule dans un fluide par une variable aléatoire à valeur dans $\mathbb{R}^3$ ... La définition d'une variable aléatoire continue fait appel à des concepts abstraits avancés que nous n'introduirons pas ici, nous nous limiterons donc aux variables aléatoires absolument continues que nous pourrons définir à partir de leur densité.

4.1 Définition d'une variable aléatoire absolument continue

Definition 4.1.1. On dira qu'une fonction f est une densité de probabilité si elle est continue (sauf éventuellement en un nombre dénombrable de points), positive et si

$$\int_{\mathbb{R}} f(t)dt = 1 \tag{4.1}$$

Definition 4.1.2. On appellera variable aléatoire continue X toute variable aléatoire à valeurs dans $\mathbb{R}$, telle que pour tout $x \in \mathbb{R}$

$$P(X \leq x) = \int_{-\infty}^{x} f(t)dt \tag{4.2}$$

où f est une densité de probabilité.

On dit que f est la densité de probabilité de X.

Proposition 4.1.3. *Si X est de densité f,*

1. *pour tout $x \in \mathbb{R}$, $P(X = x) = 0$,*

2. pour tous $x, y \in \mathbb{R}$ tels que $x \leq y$, $P(X \in [x,y]) = \int_x^y f(t)dt$.

Remarque 4.1.4. Dans la pratique pour vérifier que f est une une densité de probabilité on vérifie que :

1. f est continue (sauf éventuellement en un nombre dénombrable de points) et positive,
2. $\int_{-\infty}^{+\infty} f(t)dt = 1$.

Exemple 4.1.5.

$$f(x) = \begin{cases} 1 & si\ x \in [0,1] \\ 0 & si\ 0 \notin [0,1] \end{cases}$$

Il est clair que cette fonction est positive et continue sauf en 0 et en 1 c'est à dire en un nombre dénombrable de points donc la condition 1. de la Remarque précédente est vérifiée. La condition 2. est également vérifiée sans difficulté. f est donc bien une densité de probabilité. Si on cherche à calculer la probabilité que X soit comprise entre 1/6 et 1/3, c'est à dire $P(X \in [1/6, 1/3]) \equiv P(1/6 \leq X \leq 1/3)$ on doit effectuer le calcul suivant

$$P(X \in [1/6, 1/3]) = \int_{1/6}^{1/3} f(t)dt = \int_{1/6}^{1/3} dt = 1/6. \tag{4.3}$$

4.2 Espérance et variance

Dans ce paragraphe on rappelle des notions que l'on a déjà vues pour des variables aléatoires discrètes.

Definition 4.2.1. Soit X une variable de densité f.

1. L'espérance ou moyenne est donnée par :

$$E(X) = \int_{-\infty}^{+\infty} t f(t)dt, \tag{4.4}$$

2. Le moment d'ordre 2 est donné par :

$$E(X^2) = \int_{-\infty}^{+\infty} t^2 f(t)dt, \tag{4.5}$$

3. La variance est donné par :

$$Var(X) = E(X^2) - (E(X))^2. \tag{4.6}$$

Remarque 4.2.2. Bien entendu ces définitions sont soumises à la condition d'existence des intégrales.

Exemple 4.2.3. *Reprenons l'Exemple 4.1.4, on a :*

$$E(X) = \int_{\mathbb{R}} tf(t)dt = \int_0^1 tdt = [t^2/2]_0^1 = 1/2, \tag{4.7}$$

$$E(X^2) = \int_0^1 t^2dt = [t^3/3]_0^1 = 1/3. \tag{4.8}$$

$$\tag{4.9}$$

Rappelons maintenant les principales propriétés de l'espérance et de la variance.

Proposition 4.2.4. *Soient X et Y deux variables aléatoires continues et soit $a \in \mathbb{R}$, alors*

1. $E(aX) = aE(X)$,
2. $E(X + Y) = E(X) + E(Y)$,
3. *Si X et Y sont* ***indépendantes*** *alors* $E(XY) = E(X)E(Y)$,
4. *Si X et Y sont* ***indépendantes*** *alors* $Var(X + Y) = Var(X) + Var(Y)$

4.3 Variables aléatoires continues usuelles

Comme pour les variables aléatoires discrètes nous allons maintenant définir un certain nombre de variables aléatoires continues usuelles, en tracer la densité et en donner l'espérance et la variance.

4.3.1 La loi uniforme

La loi uniforme a déjà été introduite dans nos différents exemples, on en donne ici une définition un peu plus générale :

Definition 4.3.1. X suit une loi uniforme si et seulement si il existe a et b deux réels tels que $a < b$, et

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b] \end{cases}$$

Le cas $a = 0$ et $b = 1$ est celui que l'on a traité dans l'Exemple 4.14.

Proposition 4.3.2.

$$E(X) = \frac{b + a}{2},$$
$$E(X^2) = \frac{b^2 + ab + a^2}{3},$$
$$Var(X) = \frac{(b - a)^2}{12}.$$

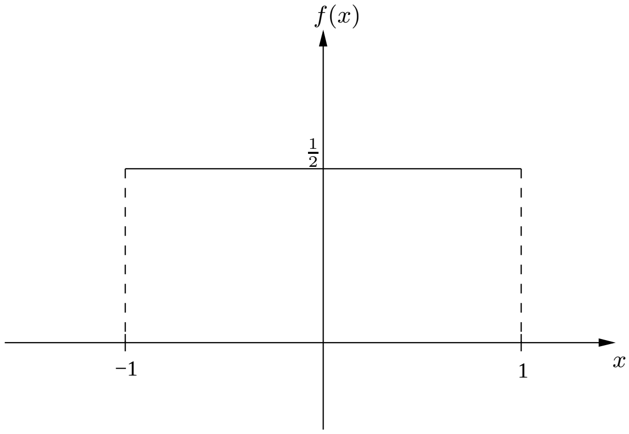
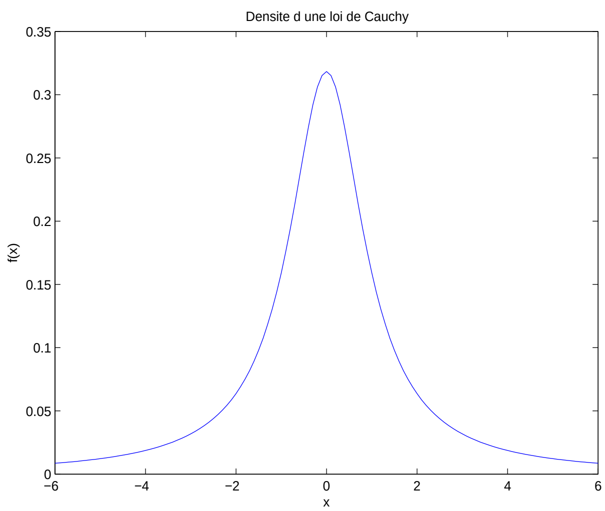


FIGURE 4.1 – Densité d'une loi uniforme, $a = -1$ et $b = 1$

4.3.2 La loi de Cauchy

Definition 4.3.3. X suit une loi de Cauchy si et seulement si pour tout $x \in \mathbb{R}$ sa densité f est donnée par

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

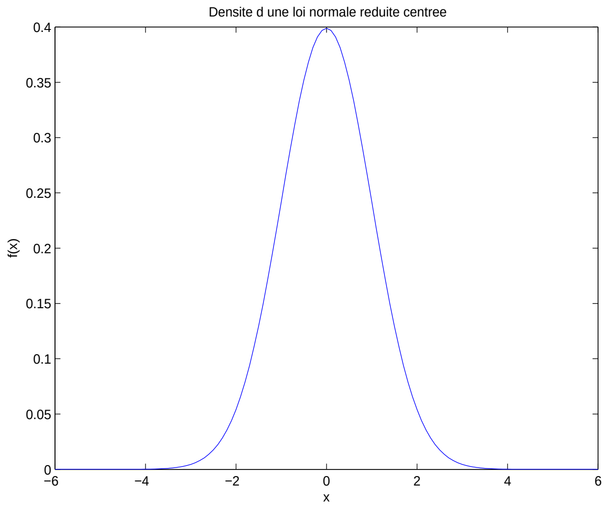


Remarque 4.3.4. La fonction $x \mapsto x/(1 + x^2)$ n'étant pas intégrable sur $\mathbb{R}$, X n'a pas d'espérance.

4.3.3 La loi Normale

Definition 4.3.5. X suit une loi Normale réduite centrée ($X \sim N(0,1)$) si et seulement si pour tout $x \in \mathbb{R}$ sa densité f est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$



Proposition 4.3.6. *Soit X une v.a. suivant une loi normale réduite centrée :*

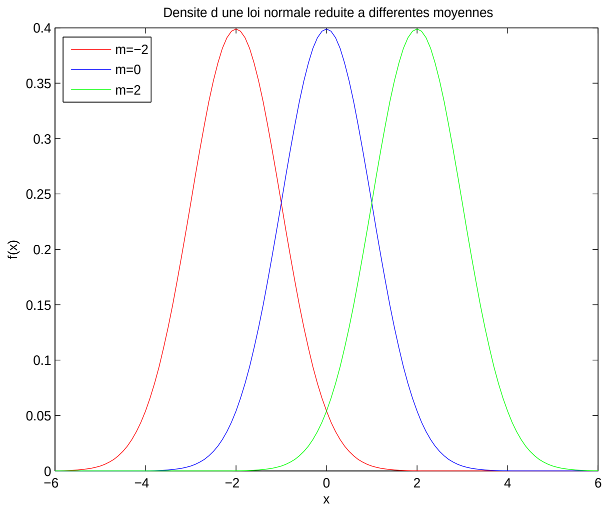
$$E(X) = 0,$$
$$E(X^2) = 1,$$
$$Var(X) = 1.$$

On définit maintenant une loi normale de moyenne m et variance σ^2 quelconque :

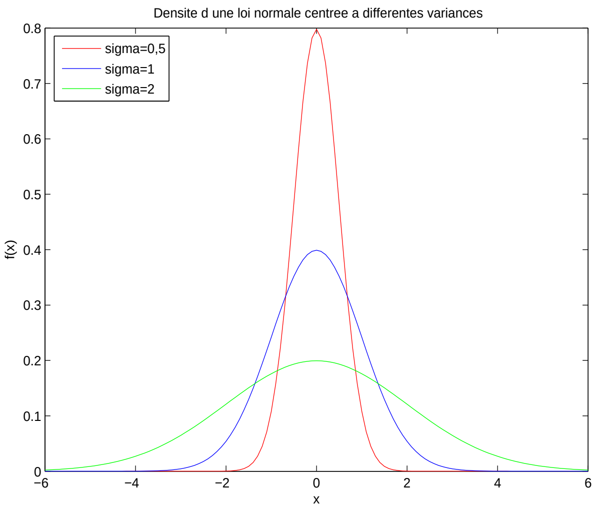
Definition 4.3.7. X suit une loi normale $X \sim N(m, \sigma)$ si et seulement si pour tout $x \in \mathbb{R}$ sa densité f est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Sur le graphe ci-dessous on a dessiné 3 densités normales de même variance mais avec trois moyennes différentes :



Sur le graphe ci-dessous on a dessiné 3 densités normales de même moyenne mais avec trois variances différentes :



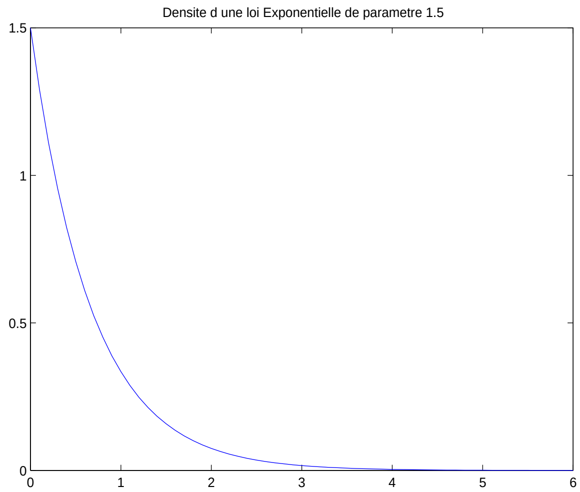
Proposition 4.3.8. *Soit X une v.a. suivant une loi normale de moyenne m et de variance σ^2 :*

$$E(X) = m,$$
$$E(X^2) = \sigma^2 + m^2,$$
$$Var(X) = \sigma^2.$$

4.3.4 La loi exponentielle

Definition 4.3.9. X suit une loi exponentielle de paramètre α ssi pour tout $x \in \mathbb{R}$ sa densité f est donnée par

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



Proposition 4.3.10. *Soit X une v.a. suivant une loi exponentielle de paramètre α :*

$$E(X) = \frac{1}{\alpha},$$
$$E(X^2) = \frac{2}{\alpha^2},$$
$$Var(X) = \frac{1}{\alpha^2}.$$

Chapitre 5

Théorèmes limites

Dans ce chapitre nous allons étudier deux résultats fondamentaux en théorie des probabilités qui sont utilisés quasi systématiquement pour des problèmes de statistiques.

5.1 Loi des grands nombres

Reprenons le lancé d'une pièce de monnaie équilibrée. Etant donné que l'on a une chance sur deux d'obtenir "face" lors d'un lancé, on a envie de dire qu'à la fin de n lancés de cette pièce on va obtenir, à peu-près, $n/2$ fois face et $n/2$ fois pile. Le but de cette première partie est de montrer ce résultat est d'interpréter le terme à "peu-près" en terme de probabilité.

5.1.1 Moyenne empirique

On commence par la définition de la **Moyenne Empirique** Soient $X_1, X_2, \cdots, X_n$, n variables aléatoires indépendantes et de même loi, On définit la somme de ces variables aléatoires

$$S_n = \sum_{i=1}^{n} X_i, \tag{5.1}$$

et la **moyenne empirique** est donnée par

$$\bar{S}_n = \frac{S_n}{n}. \tag{5.2}$$

Proposition 5.1.1. *On a :*

1. $E(\bar{S}_n) = E(X_1)$,
2. $Var(\bar{S}_n) = \frac{var(X_1)}{n}$.

Exemple 5.1.2. *On s'intéresse ici au nombre de faces obtenues après n lancés d'une pièce équilibrée. On a vu dans l'Exemple 3.3.4, que le nombre de faces obtenues S_n suit une loi binomiale de paramètres $1/2$ et n. D'après la Proposition 3.3.3 on a : $E(S_n) = n * 1/2$ et $Var(S_n) = n/4$. Par linéarité de l'espérance (Proposition 3.1.7), on a $E(\bar{S}_n) = 1/nE(S_n) =$*

1/2. Par la Proposition 3.2.4, 1. on a $Var(\bar{S}_n) = \frac{1}{n^2}Var(S_n) = 1/(4n)$. On remarque que l'on a bien $E(\bar{S}_n) = E(X_1)$ et $Var(\bar{S}_n) = Var(X_1)/n$. On remarque également que $E(S_n) = n/2$ est bien conforme à notre intuition : si on a une chance sur deux d'obtenir un certain résultat lors d'une expérience et que l'on réitère de manière indépendante cette expérience n fois, alors en moyenne on obtiendra le résultat $n/2$ fois. En fait la Loi faible des grands nombres (paragraphe 5.2.3) nous dira que l'on a plus que cela.

5.1.2 Inégalité de Markov et de Chebyshev

Ce paragraphe a pour but de donner les outils pour démontrer la loi faible des grands nombres. La preuve de ce résultat dont nous donnerons la démonstration est basée sur une inégalité très utilisée en théorie des probabilité **l'inégalité de Chebyshev**.

Proposition 5.1.3. ***Inégalité de Markov*** *Soit X une v.a. positive telle que $E(X) < +\infty$, pour tout $c > 0$ on a*

$$P(X > c) \leq \frac{E(X)}{c}. \tag{5.3}$$

Preuve
En guise d'exemple et étant donné la simplicité de la preuve de cette proposition, nous en donnons les arguments ici. On a :

$$X = X\mathbb{1}_{X\geq c} + X\mathbb{1}_{X<c}, \tag{5.4}$$

et comme X est positive,

$$X \geq X\mathbb{1}_{X\geq c}. \tag{5.5}$$

De plus $X\mathbb{1}_{X\geq c} \geq c\mathbb{1}_{X\geq c}$, on en déduit :

$$X \geq c\mathbb{1}_{X\geq c}. \tag{5.6}$$

En prenant la moyenne de chaque côté de cette inégalité on obtient la Proposition. ■
Un corollaire de la Proposition ci-dessus est :

Corollaire 5.1.4. Inégalité de Chebytshef Soit X une v.a. telle que $Var(X) < +\infty$. Pour tout $\epsilon > 0$

$$P(|X - E(X)| > \epsilon) \leq \frac{Var(X)}{\epsilon^2}. \tag{5.7}$$

5.1.3 Loi faible des grands nombres

Théorème 5.1.5. *Soit $(X_i, i \in \mathbb{N})$ une suite de variables aléatoires indépendantes de même loi telles que $Var(X_1) < \infty$. Pour tout $\epsilon > 0$,*

$$\lim_{n\to+\infty} P(|\bar{S}_n - E(X_1)| \leq \epsilon) = 1. \tag{5.8}$$

où $\bar{S}_n$ a été définie en 5.2.

Remarque 5.1.6. Ce résultat dit la chose suivante : avec une probabilité qui tend vers 1 quand n tend vers l'infini la moyenne empirique $\bar{S}_n$ des $(X_i, i \leq n)$ tend vers la moyenne d'une des variables X_i, qui ont toute la même moyenne puisque de même loi. Pour notre lancé de pièce cela signifie qu'avec une probabilité qui tend vers 1, si on fait n lancés on obtiendra bien à peu-près (à $\epsilon * n$ près, où ϵ est petit) $n/2$ fois "face".

La preuve est extrêmement simple moyennant l'inégalité de Chebytshev :
Preuve
On applique l'inégalité de Chebytshev à $X = \bar{S}_n$. En utlisant la Proposition 5.2.1 on obtient :

$$P(|\bar{S}_n - E(X_1)| > \epsilon) \leq \frac{(Var(X_1))^2}{n\epsilon^2}. \tag{5.9}$$

Par hypothèse $Var(X_1) < +\infty$, la probablité ci-dessus tend donc vers 0 quand n tend vers l'infini. Pour terminer on utilise le fait que pour tout événement A, $P(A) = 1 - P(\bar{A})$ avec $A = \{|\bar{S}_n - E(X_1)| \leq \epsilon\}$. ■

5.2 Théorème de la limite centrale

5.2.1 Le résultat

Le théorème suivant est fondamental tant en théorie des probabilités qu'en statistiques. On commence par énoncer ce théorème. Notons

$$\sigma^2 \equiv Var(X_1), \tag{5.10}$$
$$\mu \equiv E(X_1). \tag{5.11}$$

Théorème 5.2.1. *Soit $(X_i, i \in \mathbb{N})$ une suite de variables aléatoires indépendantes de même loi telles que $\sigma^2 < \infty$. Pour tout $a, b \in \mathbb{R}$, tels que $a < b$*

$$\lim_{n \to +\infty} P\left(a < \sqrt{n}\frac{\bar{S}_n - \mu}{\sigma} < b\right) = P(a < Z < b), \tag{5.12}$$

où $\bar{S}_n$ a été définie en 5.2 et $Z \sim N(0,1)$.

On peut également écrire ce résultat sous la forme :

$$\lim_{n \to +\infty} P\left(a < \frac{S_n - n\mu}{\sqrt{n}\sigma} < b\right) = P(a < Z < b). \tag{5.13}$$

Intuitivement, ce résultat montre que pour n grand, S_n suit une loi normale $N(n\mu, \sqrt{n}\sigma)$. Ce que l'on remarque est que ce résultat ne dépend pas de la loi des X_i, ce résultat est donc universel pour toute somme de variables aléatoires indépendantes de même loi.

5.2.2 Approximation d'une loi binomiale par une loi Normale ou par une loi de Poisson

Pour terminer nous donnons les deux résultats d'approximation suivants, très utilisé dans les calculs

Approximation par une loi normal

Proposition 5.2.2. *1. Soit X une variable aléatoire suivant une loi binomiale $B(n, p)$, si n suffisamment grand ($n > 100$), $np > 10$ et $n(1 - p) > 10$, alors $X \sim N(np, \sqrt{np(1 - p)})$.*

Approximation par une loi de Poisson

Proposition 5.2.3. *1. Soit X une variable aléatoire suivant une loi binomiale $B(n, p)$, si n suffisamment grand ($n > 50$) et p très petit ($np \leq 10$), alors $X \sim P(pn)$.*