Cours de Biostatistique

au profil des étudiants des Troncs Communs

Sciences de la Nature et de la Vie

Table des matières

Table des matières **1**

Préambule **4**

1 Statistiques descriptives **5**
1.1 Introduction 5
1.1.1 Concepts de base 5
1.1.2 Types de caractères 6
1.2 Tableaux statistiques et représentations graphiques 6
1.2.1 Tableau statistique relatif à un caractère qualitatif et sa représentation graphique 7
1.2.2 Tableaux statistiques relatifs à un caractère quantitatif et représentations graphiques 9
1.3 Fréquences relatives cumulées et effectifs cumulés 13
1.3.1 Variable statistique discrète 13
1.3.2 Variable statistique continue 15
1.4 Paramètres d'une variable statistique 15
1.4.1 Moyenne arithmétique 16
1.4.2 La médiane 17
1.4.3 Le mode 19
1.5 Paramètres de dispersion 20
1.5.1 Étendu 20
1.5.2 Variance et écart type 21
1.5.3 Les quartiles 22

2 Analyse combinatoire **25**
2.1 Introduction 25
2.1.1 Notion de répétition 25

2.1.2 Notion d'ordre . . . 25
2.1.3 Factoriel d'un entier n . . . 26
2.2 Arrangements . . . 26
2.2.1 Arrangement sans répétition . . . 26
2.2.2 Arrangement avec répétition . . . 27
2.3 Permutations . . . 27
2.3.1 Permutation sans répétition . . . 27
2.3.2 Permutation avec répétition . . . 28
2.4 Combinaisons . . . 28
2.4.1 Combinaison sans répétitions (sans remises) . . . 28
2.4.2 Combinaison avec répétitions (avec remises) . . . 29
2.4.3 Propriétés des combinaisons et binôme de Newton . . . 29

3 Rappel sur la théorie des probabilités 31
3.1 Expérience aléatoire et événement . . . 31
3.1.1 Expérience aléatoire . . . 31
3.1.2 Événement . . . 31
3.2 Relations et opérations entre les événements . . . 31
3.2.1 Inclusion . . . 31
3.2.2 Événement contraire . . . 32
3.2.3 Union (Disjonction) . . . 32
3.2.4 Intersection (Conjonction) . . . 32
3.2.5 Événements incompatibles (disjoints) . . . 33
3.2.6 Système complet d'événements . . . 33
3.3 Définition axiomatique de la probabilité . . . 33
3.4 Définition classique des probabilités . . . 35
3.5 Probabilités conditionnelles . . . 36
3.6 Formule des probabilités composées . . . 37
3.7 Événements indépendants . . . 37
3.8 Formule des probabilités totales . . . 38
3.9 Théorème de Bayes . . . 39

4 Variables aléatoires 41
4.1 Variables aléatoires discrètes . . . 41
4.1.1 Loi de probabilité d'une variable aléatoire discrète . . . 42
4.1.2 Fonction de distribution et de répartition . . . 43

4.1.3 Espérance mathématique d'une variable aléatoire discrète 44
4.1.4 Variance et écart type d'une variable aléatoire discrète 45
4.2 Variables aléatoires continues 47
4.2.1 Fonction de répartition 47
4.2.2 Densité de probabilité 48
4.2.3 Espérance mathématique et variance d'une variable aléatoire continue 48
4.2.4 Médiane et mode d'une variable aléatoire continue 51

5 Lois usuelles de probabilités 53
5.1 Lois de probabilités discrètes 53
5.1.1 Loi de Bernoulli 53
5.1.2 Loi binomiale 55
5.1.3 Loi de Poisson 57
5.1.4 Table de la loi de Poisson 58
5.1.5 L'approximation de la loi binomiale par la loi de Poisson 60
5.2 Lois de probabilités continues 61
5.2.1 Loi Normale 61
5.2.2 Loi Normale centrée réduite 63
5.2.3 L'approximation de la loi binomiale par une loi de normale 65
5.2.4 L'approximation de la loi de Poisson par une loi normale 67
5.2.5 Table de la loi normale centrée réduite 68

Bibliographie 69

Préambule

L'objectif assigné à ce cours est l'initiation des étudiants des troncs communs des sciences de la nature et de la vie aux traitements des données liées à leurs thématiques de travail *via* les biostatistiques. La biostatistique, qui est aussi connue sous le nom biométrie, est l'application des statistiques en biologie ; sachant que, la statistique est la science dont l'objet est de recueillir, de traiter et d'analyser des données issues de l'observation de phénomènes aléatoires, c'est-à-dire dans lesquels le hasard intervient.

La biostatistique nous permet de décrire une population donnée, selon ses attributs et ses qualités, de mesurer la précision d'une estimation ou de définir le degré d'association entre une série de caractères et d'événements. Elle englobe :

- La conception d'expériences biologiques ;
- La collecte d'informations ;
- L'analyse des données chiffrées ;
- L'interprétation des résultats et conclusion.

Ce document permet à l'étudiant de voir différents exemples d'application de la biostatistique dans les sciences expérimentales, et lui permettre de passer du stade d'observation vers le stade de description et de calculs statistiques.

Le polycopié est structuré en cinq chapitres, dont le premier aborde la statistique descriptive, qui est un ensemble d'outils permettant de décrire et d'analyser des phénomènes susceptibles d'être dénombrés et classés, elle a pour but de décrire et non d'expliquer. Le deuxième chapitre est consacré à l'introduction des méthodes de dénombrements d'objets statistiques (analyse combinatoire) utiles en théorie des probabilités, qui sera abordée dans le chapitre suivant. Les variables aléatoires discrètes et continues, ainsi que leurs paramètres de positions et de dispersions ont été abordés dans le chapitre quatre ; suivi de quelques exemples de loi de distributions dans le cas discret (loi de Bernoulli, loi Binomiale, loi de Poisson) et dans le cas continu (loi Normale, loi Normale rentrée réduite) dans le chapitre suivant.

Chapitre 1

Statistiques descriptives

1.1 Introduction

La statistique descriptive est un ensemble de méthodes permettant de décrire et d'analyser des phénomènes susceptibles d'être dénombrés et classés. Elle a pour but de décrire et non d'expliquer.

1.1.1 Concepts de base

Les observations constituent la source des informations statistiques. Avant de débuter l'étude, il faut définir l'ensemble étudié et les critères de la description chiffrée.

1. Les ensembles étudiés sont appelés **population**.
2. Les éléments de la population sont appelés **individus ou unités statistiques**.
3. Un sous ensemble de la population est **un échantillon** et sa taille correspond à son cardinal.
4. Les critères étudiés constituent des caractères ; et un caractère permet de déterminer une partition de la population.

Exemple 1.1. Nous résumons les différents concepts dans cet exemple :

Population : l'ensemble des tous les employés d'une usine.

Individu : chaque employé de l'usine.

Caractère : le salaire, l'état matrimonial, le nombre d'enfants,... etc.

Les modalités du caractère : marié, célibataire, divorcé et veuf sont les les modalités de l'état matrimonial, par exemple.

1.1.2 Types de caractères

On distingue deux types de caractères : qualitatif et quantitatif.

Caractère qualitatif

Définition 1.1. Un caractère est dit qualitatif lorsque ses modalités ne sont pas mesurables.

Exemple 1.2. Les couleurs du pelage : l'ensemble des modalités est

$$\{\text{noir}, \text{marron}, \text{blanc}, \ldots\}.$$

Caractère quantitatif

Définition 1.2. Un caractère est dit quantitatif lorsque ses modalités sont des nombres. On lui donne souvent le nom de **variable statistique**.

Une variable statistique peut être :

Discrète : Si elle prend des valeurs isolées.

Exemple 1.3. Le nombre d'enfants d'une famille.

Continue : Lorsqu'elle peut prendre n'importe quelle valeur dans son domaine de variation.

Remarque 1.1. Dans le cas continu, le nombre de ces valeurs est toujours très grand. Dans ce cas, on regroupe toutes ces valeurs en classes.
En général, toutes les grandeurs liées à l'espace (longueur, surface, volume, ...), au temps (age), à la masse (poids, teneur, ...) ou à des combinaisons (vitesse, débit, ...) sont des variables statistiques continues.

1.2 Tableaux statistiques et représentations graphiques

Soit une population composée de n individus, sur laquelle on a étudié un caractère possédant k valeurs possibles. Ces valeurs x_1, x_2, ..., x_k sont des modalités (cas qualitatif) ou des nombres (cas quantitatif).

Soient :

$$\begin{array}{lll} n_1 & \text{le nombre d'individus ayant pris la valeur} & x_1 \\ n_2 & \text{le nombre d'individus ayant pris la valeur} & x_2 \\ \vdots & & \\ n_k & \text{le nombre d'individus ayant pris la valeur} & x_k \end{array}$$

n_i est appelé **fréquence** ou **effectif** de la valeur x_i et n est l'effectif total.

On appelle **fréquence relative** ou **effectif relatif** de la valeur x_i la quantité :

$$f_i = \frac{n_i}{n} \text{ (ou en \% } f_i = \frac{n_i}{n} \times 100\%).$$

C'est la proportion d'individus ayant pris la valeur x_i.

Remarque 1.2.

$$\begin{aligned} \sum_{i=1}^{k} n_i &= n_1 + n_2 + \ldots + n_k = n. \\ \sum_{i=1}^{k} f_i &= f_1 + f_2 + \ldots + f_k = \frac{n_1}{n} + \frac{n_2}{n} + \ldots + \frac{n_k}{n} = 1. \end{aligned}$$

Modalités x_i	Effectifs n_i	Fréquences relatives $f_i = \frac{n_i}{n}$
x_1	n_1	$f_1 = \frac{n_1}{n}$
x_2	n_2	$f_2 = \frac{n_2}{n}$
$\vdots$	$\vdots$	$\vdots$
x_k	n_k	$f_k = \frac{n_k}{n}$
Total	$\sum_{i=1}^{k} n_i = n$	$\sum_{i=1}^{k} f_i = 1$

Tableau des effectifs et des fréquences relatives

1.2.1 Tableau statistique relatif à un caractère qualitatif et sa représentation graphique

Exemple 1.4. On veut étudier les lois de Mendel sur le caractère couleur de la fleur de Balsamine. Pour cela on étudiera le croisement des plantes hétérozygotes. On obtient quatre couleurs : pourpre, rose, blanc-lavande et blanche.

Population : les plantes de Balsamine.

Individu : une plante.

Caractère étudié : couleur de la fleur.

Modalités x_i	Effectif n_i	Fréquences relatives f_i	f_i en %
Pourpre	1790	0.5778	57.78 %
Rose	547	0.1766	17.66 %
Blanc-Lavande	548	0.1769	17.69 %
Blanche	213	0.0688	6.88%
Total	3098	1	100 %

Représentation graphique

L'information résumée dans un tableau statistique se traduit par un graphique pour en réaliser une synthèse visuelle.

a) Représentation par tuyaux d'orgue (diagramme en colonnes)

Dans ce cas, le graphe s'obtient en construisant autant de colonnes que des modalités du caractère qualitatif. Ces colonnes sont des rectangles de bases constantes et de hauteurs proportionnelles aux fréquences relatives.

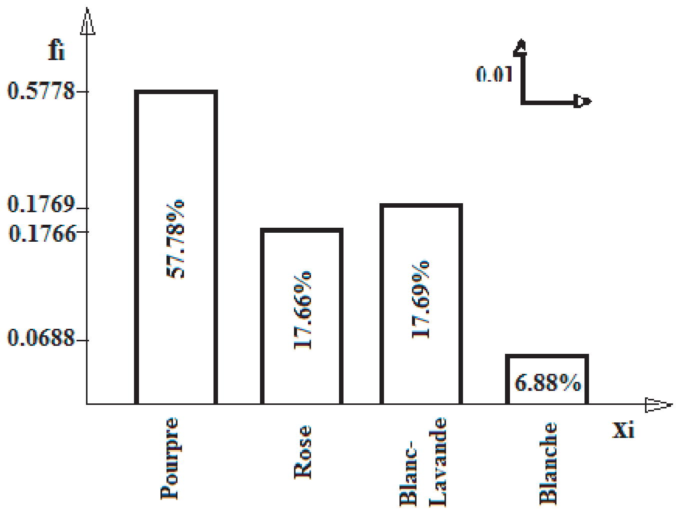


FIGURE 1.1 – Représentation en tuyaux d'orgue des fréquences relatives.

b) Représentation par le diagramme circulaire (camembert)

$$\theta_i = 360° \times f_i = 360° \times \frac{n_i}{n}.$$

Les angles correspondant de l'exemple sont :

$$\begin{aligned}
\theta_1 &= 360° \times 0.5778 = 208.01° \longrightarrow \text{Pourpre;}\\
\theta_2 &= 360° \times 0.1766 = 63.58° \longrightarrow \text{Rose;}\\
\theta_3 &= 360° \times 0.1769 = 63.68° \longrightarrow \text{Blanc-Lavande;}\\
\theta_4 &= 360° \times 0.0688 = 24.77° \longrightarrow \text{Blanche.}
\end{aligned}$$

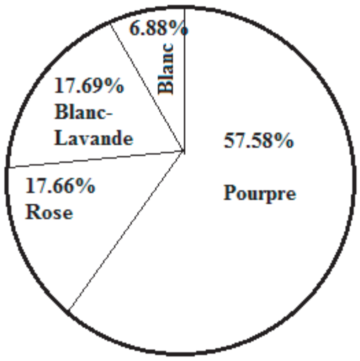


FIGURE 1.2 – Diagramme en camembert des fréquences relatives.

1.2.2 Tableaux statistiques relatifs à un caractère quantitatif et représentations graphiques

1) Cas d'une variable statistique discrète :

Exemple 1.5. Lors d'un contrôle d'une chaîne de médicaments, on s'intéresse au nombre de comprimés défectueux dans un lot. L'étude de 200 lots a donné les résultats suivants :
75 lots ont 0 comprimés défectueux ;
53 lots ont 1 comprimé défectueux ;
39 lots ont 2 comprimés défectueux ;
23 lots ont 3 comprimés défectueux ;
9 lots ont 4 comprimés défectueux ;
1 lot a 5 comprimés défectueux.

Population : l'ensemble des lots des médicaments.

Individu : un lot.

Caractère étudié : nombre de comprimés défectueux.

Modalités : 0, 1, 2, 3, 4 et 5.

Les fréquences relatives obtenues sont données dans le tableau suivant :

Modalités (Nbre de comprimés défectueux)	Nbre de lots n_i	Fréq. rel. $f_i = \frac{n_i}{n}$	Fréq. rel. en %
0	75	0.375	37.5 %
1	53	0.265	26.5 %
2	39	0.195	19.5 %
3	23	0.115	11.5 %
4	9	0.045	4.5 %
5	1	0.005	0.5 %
Total	200	1	100%

Représentation graphique :

On utilise le diagramme en batons pour représenter les effectifs n_i et les fréquences relatives f_i. Dans le cas du graphe des fréquences relatives, en joignant les sommets des batons on obtient le polygone des fréquences relatives.

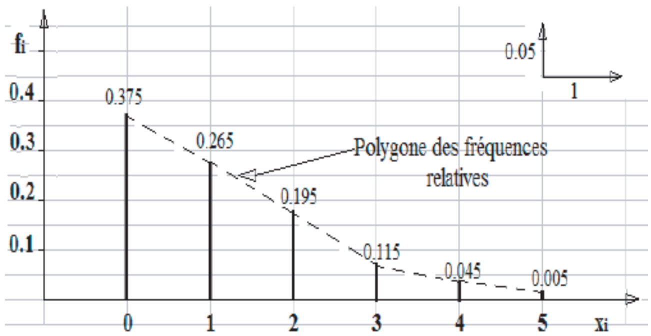


FIGURE 1.3 – Diagramme en escalier des fréquences relatives.

2) Cas d'une variable statistique continue :

Lorsque la variable statistique est continue les données sont regroupées en classes

$$[e_0, e_1[,\ [e_1, e_2[,\ [e_2, e_3[,\ ...,\ [e_{k-1}, e_k[.$$

Les modalités x_i représentent les centres c_i des classes $[e_{i-1}, e_i[$, avec :

- $c_i = \frac{e_{i-1}+e_i}{2}$, $i \in \{1, 2, ..., k\}$;
- e_{i-1} est appelé l'extrémité inférieure de la classe $[e_{i-1}, e_i[$;
- e_i est appelé l'extrémité supérieure de la classe $[e_{i-1}, e_i[$;
- $a_i = e_i - e_{i-1}$ est l'amplitude de la classe $[e_{i-1}, e_i[$;
- $e_k - e_0$ est appelé l'étendu de la variable statistique.

Exemple 1.6. Une étude faite sur la taille d'un groupe d'étudiants (en mètre) a donné les résultats suivants :

Classes	Centre c_i	Effectif n_i	Fréq. rel. f_i	(f_i en %)
$[1.5; 1.6[$	1.55	8	0.08	(8%)
$[1.6; 1.7[$	1.65	33	0.33	(30%)
$[1.7; 1.8[$	1.75	31	0.31	(31%)
$[1.8; 1.9[$	1.85	22	0.22	(22%
$[1.9; 2[$	1.95	6	0.06	(6%
Total		100	1	(100%)

Population : les étudiants du groupe.

Individu : un étudiant.

Caractère étudié : la taille d'un étudiant.

Représentation graphique :

Dans le cas d'une variable statistique continue on utilise l'histogramme.

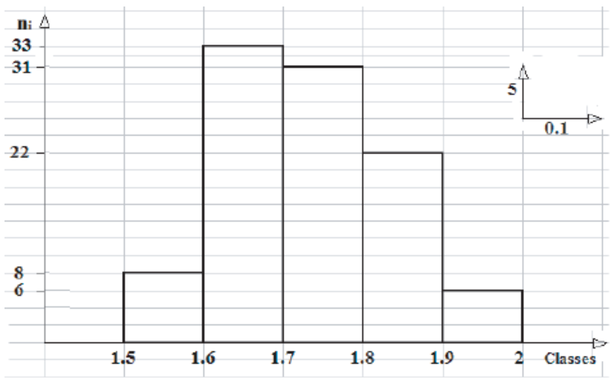


FIGURE 1.4 – L'histogramme des effectifs de l'exemple 1.6.

L'histogramme dans le cas des amplitudes inégales :

Dans ce cas les classes ont des amplitudes différentes. Du coup, il faut effectuer des corrections pour tenir compte des différences d'amplitude. Il convient de diviser les fréquences par leurs amplitudes correspondantes et on obtient ainsi, **l'amplitude corrigée** (h_i).

Exemple 1.7. Supposons que l'on regroupe les données de l'exemple précédent en classe d'amplitudes inégales.

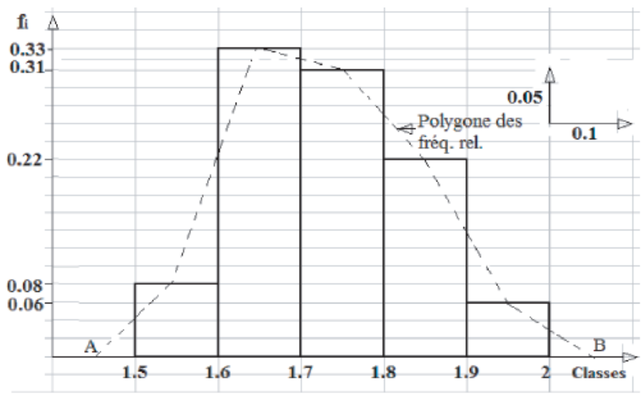


FIGURE 1.5 – L'histogramme des fréquences relatives de l'exemple 1.6.

Classes	Amplitude de la classe a_i	Effectif n_i	Fréq. rel. f_i f_i	Amplitude corrigée $h_i = \frac{f_i}{a_i}$
$[1.5; 1.7[$	0.2	41	0.41	2.05
$[1.7; 1.8[$	0.1	31	0.31	3.1
$[1.8; 2[$	0.2	28	0.28	1.4

Représentation graphique :

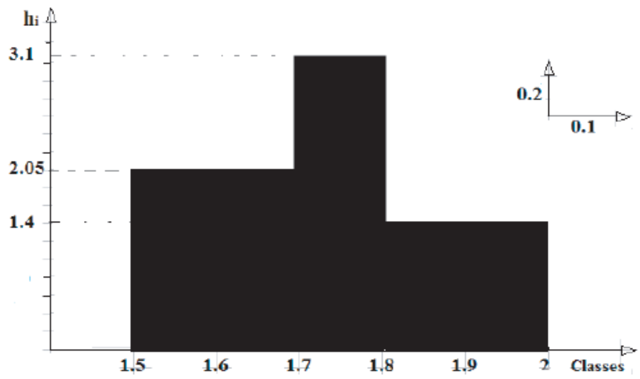


FIGURE 1.6 – Histogramme avec amplitudes inégales.

1.3 Fréquences relatives cumulées et effectifs cumulés

La fréquence relative cumulée F_i est la somme des fréquences relatives correspondantes aux valeurs de la variable statistique inférieure à x_{i+1}.

$$\begin{aligned} F_1 &= f_1; \\ F_2 &= f_1 + f_2; \\ F_3 &= f_1 + f_2 + f_3; \\ &\vdots \\ F_i &= f_1 + f_2 + \ldots + f_j + \ldots + f_i. \end{aligned}$$

La fréquence relative cumulée F_i indique la proportion des individus pour lesquels la variable statistique est inférieure à x_{i+1}. De la même façon on définit les effectifs cumulés

$$N_i = \sum_{j=1}^{i} n_j.$$

D'une manière équivalente, la fréquence relative cumulée est donnée par :

$$F_i = \frac{N_i}{n}.$$

1.3.1 Variable statistique discrète

Exemple 1.8. On reprend l'exemple du cas discret (l'exemple 1.5 des comprimés défectueux).

moins de x_i	N_i	F_i
moins de 0	0	0
moins de 1	$0 + 75 = 75$	$0 + f_1 = 0.375$
moins de 2	$75 + 53 = 128$	$f_1 + f_2 = 0.64$
moins de 3	167	0.835
moins de 4	190	0.95
moins de 5	199	0.995
moins de x_k ($x_k > 5$)	$199 + 1 = 200$	1

Remarque 1.3.

$$\sum_{i=1}^{k} f_i = 1;$$

$$F_i = 0 \text{ si } x_i < x_1;$$

$$F_i = 1 \text{ si } x_i > x_k,$$

où x_1 est la plus petite valeur observée et x_k est la plus grande valeur observée.

Soit X une variable statistique discrète et $x_1, x_2, ..., x_k$ les valeurs rangées dans l'ordre croissant. La fonction de répartition d'une v.s. discrète est définie de $\mathbb{R}$ dans $[0, 1]$ et est donnée par :

$$F(x) = \begin{cases} 0, & \text{si } x < x_1\,; \\ f_1, & \text{si } x_1 \leq x < x_2\,; \\ f_1 + f_2, & \text{si } x_2 \leq x < x_3\,; \\ \vdots & \\ f_1 + f_2 + \ldots + f_i, & \text{si } x_i \leq x < x_{i+1}\,; \\ \vdots & \\ 1, & \text{si } x \geq x_k. \end{cases}$$

Exemple 1.9. Écrivons la fonction de répartition de la variable statistique X de l'exemple des comprimés défectueux (l'exemple 1.5) :

$$F(x) = \begin{cases} 0, & \text{si } x < 0\,; \\ 0.375, & \text{si } 0 \leq x < 1\,; \\ 0.64, & \text{si } 1 \leq x < 2\,; \\ 0.835, & \text{si } 2 \leq x < 3\,; \\ 0.95, & \text{si } 3 \leq x < 4\,; \\ 0.995, & \text{si } 4 \leq x < 5\,; \\ 1, & \text{si } x \geq 5. \end{cases}$$

La courbe cumulative est la représentation graphique des fréquences relatives cumulées. Dans le cas discret, la courbe cumulative est une courbe en escalier (voir figure 1.7), dont les paliers horizontaux ont pour coordonnées (x_i, F_i).

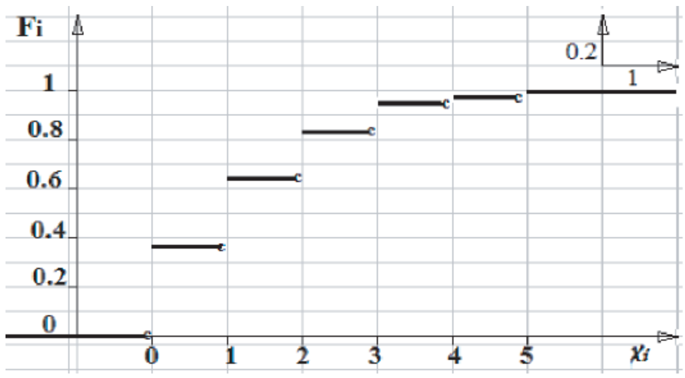


FIGURE 1.7 – Courbe des fréquences relatives cumulées (courbe cumulative).

1.3.2 Variable statistique continue

Exemple 1.10. On reprend l'exemple précédent sur la taille des étudiants. Les effectifs et les fréquences relatives cumulées sont données dans le tableau suivant :

moins de x_i	Effectifs cumulés N_i	Fréq. relatives $F_i = \frac{N_i}{n}$	cumulées (en %)
moins de 1.5	0	0	(0 %)
moins de 1.6	8	0.08	(8 %)
moins de 1.7	41	0.41	(41 %)
moins de 1.8	72	0.72	(72 %)
moins de 1.9	94	0.94	(94 %)
moins de x ($x \geq 2$)	100	1	(100 %)

La courbe cumulative est donnée dans la figure suivante :

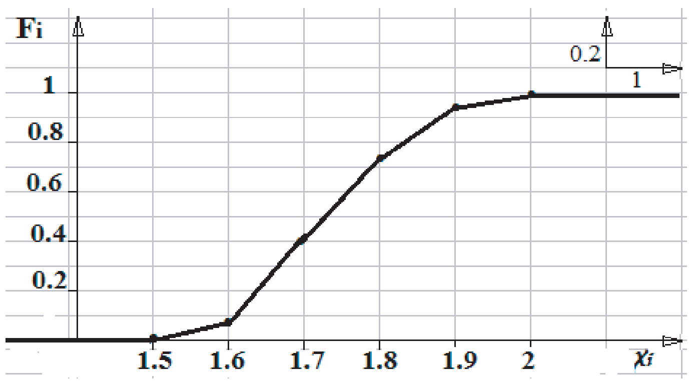


FIGURE 1.8 – Courbe des fréquences relatives cumulées du cas continu.

1.4 Paramètres d'une variable statistique

Lorsqu'on observe une représentation graphique d'une série statistique, on peut en tirer deux observations :

1. Paramètres de tendance centrale ou de position : valeurs situées au centre de la distribution statistique.
2. Paramètres de dispersion : fluctuations des observations autour de la valeur centrale, mesurées par des écarts à celles-ci.

1.4.1 Moyenne arithmétique

La moyenne arithmétique est la somme de toutes les valeurs observées divisée par le nombre total des observations.

(a) Cas d'une variable statistique discrète (données non groupées) :

Soient X une variable statistique discrète et $x_1, x_2, \ldots, x_k$ ses valeurs, pour lesquelles correspondent les effectifs $n_1, n_2, \ldots, n_k$; avec $n = \sum\limits_{i=1}^{k} n_i$ l'effectif total.

La moyenne arithmétique notée $\overline{x}$ de cette série statistique, est définie par :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i.$$

Remarque 1.4.

$$\begin{aligned}
\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i &= \frac{n_1x_1 + n_2x_2 + \ldots + n_kx_k}{n} \\
&= \frac{n_1}{n}x_1 + \frac{n_2}{n}x_2 + \ldots + \frac{n_k}{n}x_k \\
&= f_1x_1 + f_2x_2 + \ldots + f_kx_k \\
&= \sum_{i=1}^{k} f_i x_i,
\end{aligned}$$

où f_i est la fréquence relative.

Exemple 1.11. La moyenne arithmétique de 200 lots de comprimés de l'exemple 1.5 est :

$$\begin{aligned}
\overline{x} &= \frac{1}{n}\sum_{i=1}^{k} n_i x_i \\
&= \frac{75.0 + 53.1 + 39.2 + 23.3 + 9.4 + 1.5}{200} = \frac{241}{200} = 1.205.
\end{aligned}$$

(b) Cas d'une variable statistique continue (données groupées) :

Dans le cas d'une variable statistique continue, les observations sont groupées dans des classes ; et nous avons la même formule que le cas discret, sauf qu'on remplace les x_i par les centres des classes x_i :

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i = \sum_{i=1}^{k} f_i c_i.$$

Exemple 1.12. La moyenne arithmétique de la taille des étudiants de l'exemple 1.6 est :

$$\begin{aligned}\overline{x} &= \frac{1}{n}\sum_{i=1}^{k} n_i c_i \\ &= \frac{(8\times 1.55)+(33\times 1.65)+(31\times 1.75)+(22\times 1.85)+(6\times 1.95)}{100} \\ &= \frac{173.5}{100} = 1.735.\end{aligned}$$

(c) Propriétés de la moyenne arithmétique :

La moyenne de la somme des écarts d'un ensemble d'observations à leurs moyenne arithmétique est nulle $\frac{1}{n}\sum\limits_{i=1}^{k} n_i(x_i-\overline{x})=0$.

En effet,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{k} n_i(x_i-\overline{x}) &= \frac{1}{n}\sum_{i=1}^{k}(n_i x_i - n_i\overline{x}) = \frac{1}{n}\sum_{i=1}^{k} n_i x_i - \frac{1}{n}\sum_{i=1}^{k} n_i\overline{x} \\ &= \frac{1}{n}\sum_{i=1}^{k} n_i x_i - \overline{x}\frac{1}{n}\sum_{i=1}^{k} n_i = \frac{1}{n}\sum_{i=1}^{k} n_i x_i - \overline{x}\frac{n}{n} \\ &= \frac{1}{n}\sum_{i=1}^{k} n_i x_i - \overline{x} = \overline{x}-\overline{x} = 0.\end{aligned}$$

1.4.2 La médiane

La médiane est la valeur pour laquelle, on a le même nombre d'individus à gauche et à droite dans un échantillon. Elle correspond au milieu de la distribution.

(a) Cas d'une variable statistique discrète :

Pour déterminer la médiane d'un échantillon ou d'une population :

1. on classe les individus par ordre croissant ;
2. on prend celui du milieu.

Remarque 1.5. la médiane Me est la valeur qui se trouve au centre de la série statistique :

• Si la valeur de l'effectif total n est impaire, i.e. $n = 2p+1$, alors la médiane Me est la valeur qui se trouve à l'ordre $p+1$ (ou $\frac{n+1}{2}$) :

$$Me = x_{p+1} = x_{\frac{n+1}{2}}.$$

- Si la valeur de l'effectif total n est paire, i.e. $n = 2p$, alors la médiane Me est la moyenne des valeurs qui se trouve à l'ordre p et $p+1$ (ou $\frac{n}{2}$ et $\frac{n}{2}+1$) :

$$Me = \frac{x_p + x_{p+1}}{2} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}.$$

Exemple 1.13. Soit un échantillon de 11 personnes dont le poids en kg est :

$$45, 68, 89, 74, 55, 62, 56, 74, 49, 52, 63.$$

Les poids classés par ordre croissant sont :

$$\underbrace{45, 68, 49, 52, 55, 56}_{5}, \underbrace{62}_{x_6 = Me}, \underbrace{63, 68, 74, 74, 89}_{5}.$$

Si le nombre d'individus est pair, $n = 12$, on prend la moyenne entre les deux valeurs centrales.

$$\underbrace{45, 68, 49, 52, 55, 55, 56}_{6}, \underbrace{62, 63, 68, 74, 74, 89}_{6}.$$

$$45, 68, 49, 52, 55, 55, \underbrace{56, 62}, 63, 68, 74, 74, 89.$$

La médiane $Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{\frac{12}{2}} + x_{\frac{12}{2}+1}}{2} = \frac{x_6 + x_7}{2} = \frac{56+62}{2} = 59$ kg.

(b) Cas d'une variable statistique continue :

Dans ce cas la médiane est donnée par la formule suivante :

$$Me = L_i + \left(\frac{\frac{n}{2} - \sum_{i=1}^{<Me} n_i}{n_{Me}}\right).a$$

- L_i : borne inférieure de la classe médiane (classe qui divise l'effectif en deux) ;
- n : effectif total ;
- $\sum_{i=1}^{<Me} n_i$: somme des effectifs correspondant à toutes les classes inférieures à la classe médiane ;
- n_{Me} : effectif de la classe médiane ;
- a : amplitude de la classe médiane.

Exemple 1.14. Prenons le tableau des fréquences relatives cumulées de l'exemple 1.6. Pour déterminer la classe médiane, il suffit de tirer une ligne horizontale partant du point 0.5 (50%) de l'axe des fréquences relatives cumulées dans la courbe cumulative, arriver à l'ogive on descend une ligne verticale jusqu'à l'axe des x, et la classe où se situe le point d'intersection est la classe médiane. La classe médiane correspond, aussi, à la classe où les effectifs cumulées atteint ou dépasse pour la première fois le 50%.
Dans notre exemple, la classe médiane est la classe $[1.7; 1.8[$.

- $L_i = 1.7$;
- $n = 100$;
- $\sum\limits_{i=1}^{<Me} n_i = 33 + 8$;
- $n_{Me} = 31$;
- $a = 0.1$;

$$Me = L_i + (\frac{\frac{n}{2} - \sum\limits_{i=1}^{<Me} n_i}{n_{Me}}).a = 1.7 + (\frac{\frac{100}{2} - (33+8)}{31})0.1 = 1.7290.$$

1.4.3 Le mode

(a) Cas d'une variable statistique discrète :

Le mode Mo d'un ensemble d'observations est l'observation que l'on rencontre le plus fréquemment et il correspond à la modalité x_i ayant le plus grand effectif.

Remarque 1.6. Une distribution observée peut avoir plusieurs modes. Lorsqu'une distribution observée possède un seul mode, on parle de distribution unimodale. Lorsqu'une distribution observée possède deux modes, on parle de distribution bimodale.

Exemple 1.15.

1. Dans l'exemple des comprimés défectueux (exemple 1.5), le mode est 0.
2. Dans l'exemple des observations suivantes : 2, 2, 3, 5, 6, 6, 7, 8, 9, 7, 10 ; y a trois modes : 2, 6 et 7.

(b) Cas d'une variable statistique continue :

Dans le cas des données groupées en classes, le mode se calcule par la formule :

$$Mo = L_i + (\frac{\Delta_1}{\Delta_1 + \Delta_2}).a$$

- L_i : borne inférieure de la classe modale (classe correspondant au plus grand effectif) ;
- Δ_1 : excédent de l'effectif de la classe modale par rapport à l'effectif de la classe précédente ;
- Δ_2 : excédent de l'effectif de la classe modale par rapport à l'effectif de la classe suivante ;
- a : amplitude de la classe modale.

Exemple 1.16. Prenons les données de l'exemple du cas quantitatif continu (exemple 1.6) et calculons le mode.

- La classe modale est $[1.6; 1.7[$;
- $L_i = 1.6$;
- $\Delta_1 = 33 - 8$;
- $\Delta_2 = 33 - 31$;
- $a = 0.1$.

Le mode est

$$Mo = L_i + (\frac{\Delta_1}{\Delta_1 + \Delta_2}).a = 1.6 + (\frac{33-8}{(33-8)+(33-31)}) \times 0.1 = 1.6926 \text{ mètres.}$$

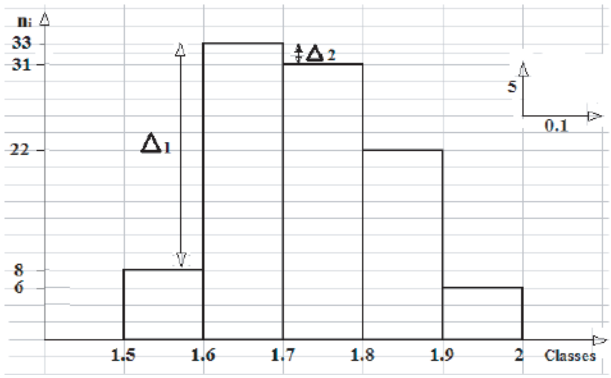


FIGURE 1.9 – L'histogramme des effectifs avec une illustration de Δ_1 et Δ_2.

1.5 Paramètres de dispersion

Les deux ensembles d'observations suivants :

$$X = \{6, 6, 7, 7, \underbrace{8}_{\overline{x}=Me}, 9, 9, 10, 10\} \text{ et } Y = \{1, 2, 4, 6, \underbrace{8}_{\overline{y}=Me}, 10, 12, 14, 15\}$$

ont la même moyenne et la même médiane $\overline{x} = \overline{y} = Me = 8$, mais ils sont différents. Le premier ensemble est moins dispersé que le deuxième.

1.5.1 Étendu

On appelle étendu, notée e, la différence entre la plus grande valeur et la plus petite valeur observée.

Exemple 1.17. L'étendu de X est $e = 10 - 6 = 4$ et l'étendu de Y est $e = 15 - 1 = 14$.

1.5.2 Variance et écart type

- **Variance**

La variance notée $V(X)$ est la moyenne des carrés des écarts des observations à la moyenne. Elle est définie par :

$$V(X) = \frac{1}{n}\sum_{1}^{k} n_i(x_i - \overline{x})^2 = \sum_{1}^{k} f_i(x_i - \overline{x})^2.$$

Propriétés

1) $V(X) \geq 0$;
2) $V(X) = \frac{1}{n}\sum\limits_{1}^{k} n_i x_i^2 - \overline{x}^2 = \sum\limits_{1}^{k} f_i x_i^2 - \overline{x}^2$.
3) Lorsqu'on compare les observations de deux variables statistiques X et Y, celle qui possède l'écart-type le plus élevé est la plus dispersée.

- **Écart type**

L'écart-type noté σ_X est la racine carrée de $V(X)$:

$$\sigma_X = \sqrt{V(X)} = \sqrt{\sum_{1}^{k} f_i(x_i - \overline{x})^2}.$$

Exemple 1.18. Cas quantitatif discret
La variance des 200 lots de médicaments (exemple 1.5)) est :

$$\begin{aligned} V(X) &= \frac{1}{n}\sum_{1}^{k} n_i x_i^2 - \overline{x}^2 \\ &= \frac{(75\times 0^2) + (53\times 1^2) + (39\times 2^2) + (23\times 3^2) + (9\times 4^2) + (1\times 5^2)}{200} - (1.205)^2 \\ &= 1.473 \end{aligned}$$

et l'écart-type $\sigma_X = \sqrt{V(X)} = \sqrt{1.473} = 1.214$.

Exemple 1.19. Cas quantitatif continu
La variance de l'exemple 1.6 est :

$$\begin{aligned} V(X) &= \frac{1}{n}\sum_{1}^{k} n_i c_i^2 - \overline{x}^2 \\ &= \frac{(8\times 1.55^2) + (33\times 1.65^2) + (31\times 1.75^2) + (22\times 1.85^2) + (6\times 1.95^2)}{100} - (1.735)^2 \\ &= 0.010875. \end{aligned}$$

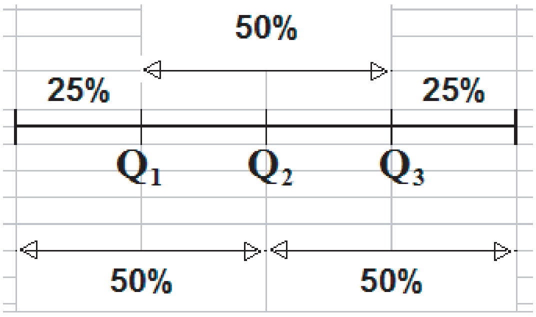
L'écart-type $\sigma_X = \sqrt{V(X)} = \sqrt{0.010875} = 0.1043$.

1.5.3 Les quartiles

On utilise couramment les quartiles Q_1, Q_2 et Q_3.
Q_1 est le quartile d'ordre $\frac{1}{4}$, représente 25% de l'échantillon ;
Q_2 est le quartile d'ordre $\frac{1}{2}$, représente 50% de l'échantillon ;
Q_3 est le quartile d'ordre $\frac{3}{4}$, représente 75% de l'échantillon.

Intervalle interquartile

(Q_1, Q_3) contient 50% de la population laissant à droite 25% et à gauche 25%. Cet intervalle est donnée par : $Q_1 - Q_3$.
Pour déterminer l'intervalle interquartile, il faut déterminer d'abord Q_1 et Q_3.



Le premier quartile Q_1

(a) Cas discret

Q_1 est la valeur x_i dont le rang (la position) est le plus petit entier qui suit $\frac{n}{4}$.

Exemple 1.20. Dans l'exemple des observations suivantes : 2, 3, 4, 5, 6, 6, 7, 7, 8, 9, 10 ; on a :
$n = 11$ et $\frac{n}{4} = \frac{11}{4} = 2.25$. Le plus petit entier qui suit $\frac{n}{4} = 2.25$ est 3, alors Q_1 est la troisième valeur. D'où $Q_1 = x_3 = 4$.

Exemple 1.21. Reprenons l'exemple des comprimés défectueux (exemple 1.5), où on a $\frac{n}{4} = \frac{200}{4} = 50$ et la valeur pour laquelle les effectifs cumulés atteignent ou dépassent pour la première fois 50 est "moins 1", c'est à dire la valeur de 0, alors $Q_1 = 0$.

(b) Cas continu

Dans ce cas le premier quartile est donné par la formule suivante :

$$Q_1 = L_i + \left(\frac{\frac{n}{4} - \sum_{i=1}^{<Q_1} n_i}{n_{Q_1}}\right).a$$

- L_i : borne inférieure de la classe de Q_1 ;
- n : effectif total ;
- $\sum\limits_{i=1}^{<Q_1} n_i$: somme des effectifs correspondant à toutes les classes inférieures à la classe de Q_1 ;
- n_{Q_1} : effectif de la classe de Q_1 ;
- a : amplitude de la classe de Q_1.

Exemple 1.22. Prenons le tableau de l'exemple 1.6. Pour déterminer la classe de Q_1, il suffit de tirer une ligne horizontale partant du point 0.25 (25%) de l'axe des fréquences relatives cumulées dans la courbe cumulative, arriver à l'ogive on descend une ligne verticale jusqu'à l'axe des x et la classe où se situe le point d'intersection est la classe de Q_1.
La classe de Q_1 correspond, à la classe où les effectifs cumulées atteignent ou dépassent pour la première fois 25% de l'effectif total.
Dans notre exemple, la classe de Q_1 est la classe $[1.6; 1.7[$.

- $L_i = 1.6$;
- $n = 100$;
- $\sum\limits_{i=1}^{<Q_1} n_i = 8$;
- $n_{Q_1} = 33$;
- $a = 0.1$;

$$Q_1 = L_i + \left(\frac{\frac{n}{4} - \sum\limits_{i=1}^{<Q_1} n_i}{n_{Q_1}}\right).a = 1.6 + \left(\frac{\frac{100}{4} - 8}{33}\right)0.1 = 1.65515.$$

Le troisième quartile Q_3

(a) Cas discret

Q_3 est la valeur x_i dont le rang (la position) est le plus petit entier qui suit $\frac{3n}{4}$.

Exemple 1.23. Dans le même exemple des observations : 2, 3, 4, 5, 6, 6, 7, 7, 8, 9, 10 ; on a :
$n = 11$ et $\frac{3n}{4} = \frac{3\times 11}{4} = 8.25$. Le plus petit entier qui suit $\frac{3n}{4} = 8.25$ est 9, alors Q_3 est la $9^{\text{ème}}$ valeur. D'où $Q_3 = x_9 = 8$.

Exemple 1.24. Dans l'exemple des comprimés défectueux (exemple 1.5), on a $\frac{3n}{4} = \frac{3\times 200}{4} = 150$ et la valeur où les effectifs cumulés atteignent ou dépassent pour la première fois 150 est "moins 3", c'est à dire la valeur de 2, alors $Q_3 = 2$.

(b) Cas continu

Dans le cas continu, le troisième quartile est donné par la formule suivante :

$$Q_3 = L_i + (\frac{\frac{3n}{4} - \sum_{i=1}^{<Q_3} n_i}{n_{Q_3}}).a$$

- L_i : borne inférieure de la classe de Q_3 ;
- n : effectif total ;
- $\sum_{i=1}^{<Q_3} n_i$: somme des effectifs correspondant à toutes les classes inférieures à la classe de Q_3 ;
- n_{Q_3} : effectif de la classe de Q_3 ;
- a : amplitude de la classe de Q_3.

Exemple 1.25. Pour déterminer la classe de Q_3 de l'exemple 1.6, également, il suffit de tirer une ligne horizontale partant du point 0.75 (75%) de l'axe des fréquences relatives cumulées dans la courbe cumulative, arriver à l'ogive on descend une ligne verticale jusqu'à l'axe des x et la classe où se situe le point d'intersection est la classe de Q_3.
La classe de Q_3 correspond, à la classe où les effectifs cumulées atteignent ou dépassent pour la première fois 75% de l'effectif total.
Dans notre exemple, la classe de Q_3 est la classe $[1.8; 1.9[$.

- $L_i = 1.8$;
- $n = 100$;
- $\sum_{i=1}^{<Q_3} n_i = 8 + 33 + 31$;
- $n_{Q_3} = 22$;
- $a = 0.1$;

$$Q_3 = L_i + (\frac{\frac{3n}{4} - \sum_{i=1}^{<Q_3} n_i}{n_{Q_3}}).a = 1.8 + (\frac{\frac{3\times 100}{4} - (8 + 33 + 31)}{22})0.1 = 1.8136.$$

Chapitre 2

Analyse combinatoire

L'analyse combinatoire est une branche des mathématiques qui étudie comment compter les objets. Elle fournit des méthodes de dénombrements particulièrement utile en théorie des probabilités.

2.1 Introduction

L'analyse combinatoire a pour but de compter les dispositions qui peuvent être formées à partir des éléments d'un ensemble fini d'objets. Un objet est caractérisé par :

- la place qu'il occupe dans la disposition ;
- le nombre de fois où il peut apparaître.

2.1.1 Notion de répétition

Si un élément apparaît plus d'une fois dans une disposition, on dit que la disposition est avec répétition ; sinon, la disposition est dite sans répétition.

2.1.2 Notion d'ordre

Une disposition est dite ordonnée, si lorsqu'à chaque fois qu'un élément change de place (ou de position) la disposition change.

Exemple 2.1. On considère un ensemble E ayant trois éléments $E = \{a, b, c\}$. Choisir deux éléments dans cet ensemble peut se faire de plusieurs façons différentes.
Le tableau suivant, nous donne tous les cas possibles :

disposition	avec répétition	sans répétition
avec ordre (ordonnée)	aa, ab, ac, ba, bb, bc, ca, cb, cc	ab, ac, ba, bc, ca, cb
sans ordre (non ordonnée)	aa, ab, ac, bb, bc, cc	ab, ac, bc

2.1.3 Factoriel d'un entier n

Étant donné un entier n, le factoriel de n, noté $n!$ est :

$$n! = n \times (n-1) \times (n-2) \times ... \times 3 \times 2 \times 1.$$

Par convention, on a : $0! = 1$.

Exemple 2.2.

$$\begin{aligned} 5! &= 5 \times 4 \times 3 \times 2 \times 1 = 120. \\ 10! &= 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3628800. \end{aligned}$$

2.2 Arrangements

Définition 2.1. Étant donné un ensemble E de n objets, un arrangement de p de ces objets est une suite **ordonnée** de p objets pris parmi ces n objets.

On distingue deux types d'arrangements : avec et sans répétition.

2.2.1 Arrangement sans répétition

On appelle arrangement sans répétition de n objets p à p, toute disposition **ordonnée** de p objets choisis parmi les n objets **sans répétitions**.

Le nombre d'arrangements sans répétition, noté A_n^p, est :

$$A_n^p = n \times (n-1) \times (n-2)..... \times (n-p+1) = \frac{n!}{(n-p)!},$$

avec $1 \leq p \leq n$.

Dans un arrangement sans répétition, les p objets de la liste sont tous distincts. Cela correspond à un tirage **sans remise** et **avec ordre**.

Exemple 2.3. Combien de mots de trois lettres ne contenant pas plus d'une fois la même lettre peut-on former avec les lettres de l'alphabet ?

$$A_{26}^3 = \frac{26!}{(26-3)!} = 26 \times 25 \times 24 = 15600 \text{ mots}.$$

2.2.2 Arrangement avec répétition

On appelle arrangement avec répétition de n objets p à p, toute disposition **ordonnée** de p objets choisis parmi les n objets **avec répétitions**.

Le nombre d'arrangements avec répétition est :

$$n^p = \underbrace{n \times n \times n ... \times n}_{p \text{ fois}}$$

avec $1 \leq p \leq n$.

Dans un arrangement sans répétition, les p objets de la liste ne sont pas nécessairement tous distincts. Cela correspond à un tirage **avec remise** et **avec ordre**.

Exemple 2.4. Combien de mots de deux lettres peut-on former avec les lettres de l'alphabet ?

$$26^2 = 26 \times 26 = 676 \text{ mots}.$$

2.3 Permutations

Définition 2.2. Étant donné un ensemble E de n objets. On appelle permutation de n objets distincts **toute suite ordonnée** de n objets ou **tout arrangement** n à n de ces objets.

2.3.1 Permutation sans répétition

C'est le cas particulier de l'arrangement sans répétition de p objets parmi n objets, lorsque $p = n$.

Le nombre de permutations de n objets est :

$$P_n = n!.$$

Remarque 2.1.

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

Exemple 2.5. Le nombre de manières de placer huit convives (invités) autour d'une table est

$$P_8 = 8! = 40320.$$

2.3.2 Permutation avec répétition

Dans le cas où il existe k objets identiques parmi les n objets, alors

$$P_n = \frac{n!}{k!}.$$

Exemple 2.6. Le nombre de mots possibles (avec ou sans signification) que l'on peut formé en permutant les 8 lettres du mot "Quantité", est

$$P_8 = \frac{8!}{2!} = 20160 \text{ mots, car on a le } \mathbf{t} \text{ 2 fois.}$$

Et en considérant le mot "Natation", le nombre de mots possibles est

$$P_8 = \frac{8!}{2!2!2!} = 5040 \text{ mots, car on a le } \mathbf{n} \text{ 2 fois, le } \mathbf{a} \text{ 2 fois et le } \mathbf{t} \text{ 2 fois.}$$

2.4 Combinaisons

2.4.1 Combinaison sans répétitions (sans remises)

Définition 2.3. Étant donné un ensemble E de n objets. On appelle combinaisons de p objets **tout ensemble** de p objets pris parmi les n objets sans remise.

Le nombre de combinaisons de p objets parmi n et sans remise, est :

$$C_n^p = \frac{n!}{(n-p)!p!},$$

avec $1 \leq p \leq n$.

Remarque 2.2.

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

Exemple 2.7. Le tirage au hasard de 5 cartes dans un jeu de 32 cartes (main de poker) est une combinaison avec $p = 5$ et $n = 32$. Le nombre de tirages possibles est

$$C_{32}^5 = \frac{32!}{(32-5)!5!} = 409696.$$

Exemple 2.8. La formation d'une délégation de 2 étudiants parmi un groupe de 20 constitue une combinaison avec $p = 2$ et $n = 20$. Le nombre de délégations possibles est

$$C_{20}^2 = \frac{20!}{(20-2)!2!} = 190.$$

2.4.2 Combinaison avec répétitions (avec remises)

Le nombre de combinaisons de p objets parmi n et avec remise, est :

$$C_{n+p-1}^{p} = \frac{(n+p-1)!}{p!(n-1)!}.$$

Exemple 2.9. Soit la constitution de mots de 3 lettres à partir d'un alphabet à 5 lettres avec remise.
Le nombre de mots est

$$C_{5+3-1}^{3} = C_7^3 = 35.$$

On distingue 3 cas possibles :

- C_5^3 nombre de mots de 3 lettres différentes et sans ordre ;
- $2C_5^2$ nombre de mots de 2 lettres différentes et une lettre redondante ;
- C_5^1 nombre de mots de 3 lettres identiques ;

au total, on a $C_5^3 + 2C_5^2 + C_5^1 = C_7^3 = 35$ mots possibles.

2.4.3 Propriétés des combinaisons et binôme de Newton

(1)

$$C_n^0 = C_n^n = \frac{n!}{n!} = 1;$$

$\forall n \geq 1$, on a

$$C_n^1 = C_n^{n-1} = \frac{n!}{(n-1)!} = n;$$

$\forall n \geq 2$, on a

$$C_n^2 = C_n^{n-2} = \frac{n(n-1)}{2};$$

(2) Par récurrence, on déduit :
Si $0 \leq p \leq n$, on a

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

Combinaisons composées ou formule de Pascal
Si $0 \leq p \leq n-1$, on a

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

(3) **Binôme de Newton**
Le théorème du binôme de Newton donne l'expression générale du développement de

$(a+b)^n$.

$$\begin{aligned} (a+b)^n &= \sum_{p=0}^{n} C_n^p a^{n-p} b^p \\ &= C_n^0 a^n b^0 + C_n^1 a^{n-1} b^1 + \ldots + C_n^n a^0 b^n \\ &= a^n + n a^{n-1} b + \ldots + b^n. \end{aligned}$$

Exemple 2.10. Pour $n = 4$, on a

$$\begin{aligned} (a+b)^4 &= \sum_{p=0}^{4} C_4^p a^{4-p} b^p \\ &= C_4^0 a^4 b^0 + C_4^1 a^3 b^1 + \ldots + C_4^4 a^0 b^4 \\ &= a^4 + 4a^3 b + 6a^2 b^2 + 4ab^3 + b^4. \end{aligned}$$

Pour $n = 5$,

$$(a+b)^5 = a^5 + 5a^4 b + 10a^3 b^2 + 10a^2 b^3 + 5ab^4 + b^5.$$

Remarque 2.3. Lorsque $a = b = 1$, on a $(1+1)^n = 2^n$ et $2^n = \sum_{p=0}^{n} C_n^p$. Alors

$$2^n = C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^n.$$

Chapitre 3

Rappel sur la théorie des probabilités

3.1 Expérience aléatoire et événement

3.1.1 Expérience aléatoire

Une expérience aléatoire (e.a) est toute expérience dont le résultat est régi par le hasard.

Exemple 3.1. Le jet d'une pièce de monnaie et l'observation de la face supérieure est une expérience aléatoire qui conduit à deux résultats possibles : Face (F) ou Pile (P).

Définition 3.1. L'ensemble de tous les résultats possibles d'une e.a. est appelé **ensemble fondamental** et on le note généralement Ω.

Exemple 3.2. Lorsqu'on jette un dé (à six faces numérotées), si on s'intéresse au nombre obtenu sur la face supérieure, l'ensemble fondamental est $\Omega = \{1, 2, 3, 4, 5, 6\}$.

3.1.2 Événement

Un événement de Ω est un sous ensemble de Ω. Un événement peut être élémentaire (un seul élément) ou composé (plusieurs éléments).

Exemple 3.3. Lorsqu'on jette un dé à six faces numérotées $\Omega = \{1, 2, 3, 4, 5, 6\}$,
l'événement A **:** "avoir le chiffre 2", est un événement élémentaire $A = \{2\} \subset \Omega$.
l'événement B **:** "avoir un chiffre pair", est un événement composé $B = \{2, 4, 6\} \subset \Omega$.

3.2 Relations et opérations entre les événements

3.2.1 Inclusion

Soient A et B deux événements associés à une expérience aléatoire. On dira que A est inclu dans B (ou A implique B), si la réalisation de A entraîne nécessairement la réalisation de B. On le note $A \subset B$ (ou $A \Rightarrow B$).

Exemple 3.4. Dans l'exemple précédent (exemple 3.3) $A = \{2\} \subset B = \{2, 4, 6\}$, si A est réalisé alors B est réalisé.

3.2.2 Événement contraire

On appelle événement contraire de l'événement A le complémentaire de A dans Ω, noté $\overline{A}$, l'événement réalisé lorsque A n'est pas réalisé et vice versa.

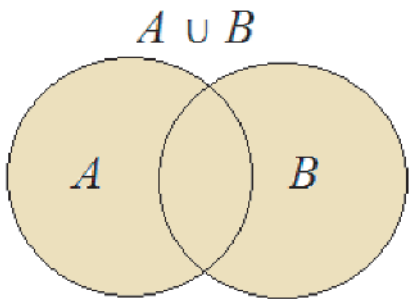
Remarque 3.1. Soit A un événement de Ω et $\overline{A}$ son événement contraire.

1. Si A est réalisé alors $\overline{A}$ n'est pas réalisé.
2. Si A n'est pas réalisé alors $\overline{A}$ est réalisé.
3. L'ensemble fondamental Ω est toujours réalisé, on l'appelle **événement certain**. Son événement contraire est **l'événement impossible**, noté $\emptyset$.

Exemple 3.5. Si on prend l'événement B dans l'exemple précédent, "avoir un chiffre pair" $B = \{2, 4, 6\}$; alors son événement contraire $\overline{B}$ est "avoir un chiffre impair" $\overline{B} = \{1, 3, 5\}$.

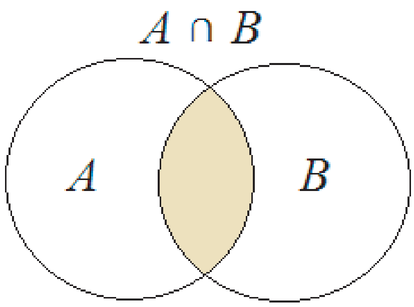
3.2.3 Union (Disjonction)

On dit que l'événement "A ou B", noté $(A \cup B)$, est réalisé si l'un au moins des deux événements est réalisé (i.e. A est réalisé **ou** B est réalisé).



3.2.4 Intersection (Conjonction)

L'événement "A et B", noté $(A \cap B)$, est réalisé lorsque A est réalisé **et** B est réalisé



3.2.5 Événements incompatibles (disjoints)

Les événements A et B sont dits incompatibles si la réalisation de l'un exclut la réalisation de l'autre. Autrement dit, si l'un est réalisé l'autre ne se réalisera pas (deux événements qui ne peuvent se réaliser à la fois) :

$$A \text{ et } B \quad \text{sont incompatibles} \Leftrightarrow A \cap B = \emptyset.$$

Exemple 3.6. B et $\overline{B}$ sont incompatibles.

3.2.6 Système complet d'événements

Soient Ω l'ensemble fondamentale associé à une expérience aléatoire et A_1, A_2,..., A_n des événements de Ω. Les événements A_1, A_2,..., A_n forment un système complet d'événements, si les conditions suivantes sont vérifiées :

1. Les A_i sont réalisables $(A_i \neq \emptyset)$, $\forall i = 1, 2, ..., n$.
2. Les A_i sont incompatibles 2 à 2 : $A_i \cap A_j = \emptyset$, $\forall (i, j) \in \{1, 2, ..., n\}$, $i \neq j$.
3. $\bigcup\limits_{i=1}^{n} A_i = \Omega$.

Exemple 3.7. Dans le jet d'un dé (une fois), on a : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Les événements $A_1 = \{1\}$, $A_2 = \{2\}$, $A_3 = \{3\}$, $A_4 = \{4\}$, $A_5 = \{5\}$ et $A_6 = \{6\}$ forment un système complet d'événements.

3.3 Définition axiomatique de la probabilité

Soient deux événements A et B dans Ω.

1. La probabilité d'un événement A est un nombre compris entre 0 et 1 :

$$0 \leq p(A) \leq 1.$$

2. La probabilité de l'événement certain Ω est égale à 1 et la probabilité de l'événement impossible $\emptyset$ est 0 :
$$p(\Omega) = 1 \text{ et } p(\emptyset) = 0.$$
3. Soient A et B deux événements incompatibles, $(A \cap B = \emptyset)$, alors
$$p(A \cup B) = p(A) + p(B).$$
4. Si $A \subseteq B$, alors $p(A) \leq p(B)$.
5. **Généralisation à n événements** : Soient A_1, A_2,..., A_n des événements incompatibles 2 à 2. Alors
$$p(A_1 \cup A_2 \cup ... \cup A_n) = p(A_1) + p(A_2) + ... + p(A_n), \text{ i.e. } p(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} p(A_i).$$

Conséquence. Soit A un événement et $\overline{A}$ son contraire, alors
$$p(\overline{A}) = 1 - p(A).$$

En effet,
$$\begin{aligned} A \cup \overline{A} = \Omega, \text{ donc } p(A \cup \overline{A}) = p(\Omega) &\Leftrightarrow p(A) + p(\overline{A}) = p(\Omega) = 1 \\ &\Leftrightarrow p(\overline{A}) = 1 - p(A). \end{aligned}$$

Remarque 3.2. Tout calcul conduisant à des valeurs de probabilités négatives ou supérieures à 1 est faux.

Exemple 3.8. Reprenons l'exemple du jet d'un dé à 6 faces équilibrées, où
$$\Omega = \{1, 2, 3, 4, 5, 6\}$$
avec les événements :
$$A_1 = \{1\}, A_2 = \{2\}, A_3 = \{3\}, A_4 = \{4\}, A_5 = \{5\} \text{ et } A_6 = \{6\}.$$
Les événements $A_i, i = 1, 2, ..., 6$ sont tous incompatibles et $p(A_i) = \frac{1}{6}$ pour tout $i = \overline{1, 6}$.
$A_1 \cup A_2 \cup ... \cup A_n = \{1, 2, 3, 4, 5, 6\} = \Omega$.
$p(A_1 \cup A_2 \cup ... \cup A_n) = p(A_1) + p(A_2) + ... + p(A_n) = 6 \times \frac{1}{6} = 1$.

3.4 Définition classique des probabilités

A chaque événement A d'une expérience aléatoire (e.a.) est associé un nombre que l'on note $p(A)$ compris entre 0 et 1 qui mesure la probabilité de la réalisation de A. Si une e.a. a N cas possibles et parmi ces N cas, il y a n cas favorables à la réalisation de l'événement A, on définit la probabilité de la réalisation de A par :

$$p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\text{nombre d'éléments de } \ A}{\text{nombre d'éléments de } \ \Omega}.$$

D'une manière équivalente

$$p(A) = \frac{n}{N}.$$

Exemple 3.9. Dans le jet d'un dé à six faces équilibrées, soit A l'événement "avoir un nombre pair".
Nombre de cas possibles est 6 : $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Nombre de cas favorables est 3 : $A = \{2, 4, 6\}$.

$$p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{3}{6}.$$

Théorème 3.1. *Soient A et B deux événements de l'espace fondamental Ω, alors :*

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

Exemple 3.10. Dans le jet du dé à 6 faces équilibrées, considérons les événements :

A : "avoir un nombre pair" ;
B : "avoir un multiple de 3".

On a

$$A = \{2, 4, 6\}, B = \{3, 6\} \text{ et } A \cap B = \{6\},$$

alors

$$p(A) = \frac{3}{6}, p(B) = \frac{2}{6} \text{ et } p(A \cap B) = \frac{1}{6}.$$

D'où

$$p(A \cup B) = p(A) + p(B) - P(A \cap B) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6}.$$

En effet,

$$A \cup B = \{2, 3, 4, 6\} \Rightarrow p(A \cup B) = \frac{4}{6}.$$

3.5 Probabilités conditionnelles

Considérons le jet de deux dés parfaits et soit A l'événement : "la somme des points obtenus est au moins égale à 10".
Les cas qui donnent au moins 10 sont

$$A = \{(4,6),(5,5),(5,6),(6,4),(6,5)\}, \text{ et } p(A) = \frac{6}{36} = \frac{1}{6}.$$

1. Supposons que le premier dé nous donne le chiffre 3 (événement B : "obtenir le chiffre 3 sur la surface supérieure du premier dé"). Alors, l'événement A est devenu irréalisable (A et B incompatibles). Nous dirons que la probabilité de A sachant que B est réalisé est nulle, et nous écrivons $p(A/B) = 0$.
2. Supposons maintenant que le premier dé amène un 6 (événement C). Pour atteindre ou dépasser 10, il faut avoir sur la face supérieure du second dé : 4, 5, ou 6. On aura 3 chance sur 6 et $p(A/B) = \frac{3}{6} = \frac{1}{2}$.

Définition 3.2. Soient A et B deux événements, tels que $p(B) \neq 0$. La probabilité conditionnelle de A sachant que B est réalisé est donnée par la formule :

$$p(A/B) = \frac{p(A \cap B)}{p(B)}.$$

Exemple 3.11. Une urne contient 2 boules rouges et 3 boules blanches. On tire une boule, on la garde, puis on tire une autre.

(i) Quelle est la probabilité de tirer une boule rouge au deuxième tirage, sachant que on a tiré une boule rouge au premier tirage ?

(ii) Quelle est la probabilité d'avoir deux boules rouges au cours des deux tirages (une boule rouge dans chaque tirage) ?

Solution

Posons les événements suivants :
A_1 : "avoir une boule rouge au premier tirage".
A_2 : "avoir une boule rouge au deuxième tirage".
$A_1 \cap A_2$: "avoir une boule rouge dans chaque tirage (deux boules rouges)".

(i) On a

$$p(A_1) = \frac{2}{5}$$

et la probabilité de tirer une boule rouge au deuxième tirage, sachant que on a tiré une boule rouge au premier tirage est :

$$p(A_2/A_1) = \frac{1}{4}.$$

(ii) Par définition

$$p(A_2/A_1) = \frac{p(A_1 \cap A_2)}{p(A_1)},$$

alors

$$p(A_1 \cap A_2) = p(A_1) \times p(A_2/A_1) = \frac{2}{5} \times \frac{1}{4} = \frac{1}{10}.$$

3.6 Formule des probabilités composées

Pour tout événement A et B tels que $p(A) \neq 0$ et $p(B) \neq 0$, on a :

$$p(A/B) = \frac{p(A \cap B)}{p(B)}, \text{ alors } p(A \cap B) = p(A/B)p(B);$$

$$p(B/A) = \frac{p(A \cap B)}{p(A)}, \text{ alors } p(A \cap B) = p(B/A)p(A).$$

Des deux formules énoncées, on déduit

$$p(A \cap B) = p(A/B)p(B) = p(B/A)p(A).$$

Exemple 3.12. Une urne contient trois boules blanches et deux boules noires. On tire deux boules successivement et sans remise.
Quelle est la probabilité pour que la première boule soit noire et que la deuxième soit blanche ?

Solution
Posons les événements suivants :
A : "tirer une boule noire au premier tirage".
B : "tirer une boule blanche au deuxième tirage".
On a

$$p(A) = \frac{2}{5} \text{ et } p(B/A) = \frac{3}{4},$$

d'où

$$p(A \cap B) = p(B/A)p(A) = \frac{3}{4} \times \frac{2}{5} = \frac{3}{10}.$$

3.7 Événements indépendants

Deux événements A et B sont dit indépendants si la réalisation ou la non réalisation de l'un ne modifie pas la réalisation ou la non réalisation de l'autre. A et B sont indépendants si

$$p(A \cap B) = p(A)p(B).$$

Conclusion

$$p(A/B) = \frac{p(A \cap B)}{p(B)} = \frac{p(A)p(B)}{p(B)} = p(A).$$

Exemple 3.13. Dans le jet d'un dé à six faces numérotées, considérons les événements :
A : "avoir un nombre pair".
B : "avoir un multiple de 3".
On a donc $A = \{2, 4, 6\}$, $B = \{3, 6\}$ et $A \cap B = \{6\}$.
Alors $p(A) = \frac{3}{6}$, $p(B) = \frac{2}{6}$ et $p(A \cap B) = \frac{1}{6}$.
On a aussi, $p(A) \times p(B) = \frac{3}{6} \times \frac{2}{6} = \frac{1}{6}$.
Donc $p(A \cap B) = p(A) \times p(B) = \frac{1}{6}$, c'est-à-dire A et B sont indépendants.

3.8 Formule des probabilités totales

Soient $\{A_1, A_2, ..., A_n\}$ une famille d'événements constituant un système complet d'événements de Ω, c'est-à-dire :

$$A_i \neq \emptyset, \forall i = \overline{1, n};\ A_i \cap A_j = \emptyset, \forall i \neq j \text{ et } \bigcup_{i=1}^{n} A_i = \Omega.$$

Soit B un événement quelconque de Ω, alors

$$p(B) = p(B/A_1)p(A_1) + p(B/A_2)p(A_2) + ... + p(B/A_n)p(A_n),$$

identiquement équivalent à

$$p(B) = \sum_{i=1}^{n} p(B/A_i)p(A_i).$$

Exemple 3.14. Trois machines A, B et C produisent respectivement 40%, 35% et 25% du nombre total de comprimés fabriquées par un laboratoire pharmaceutique. Chacune de ces machines produit respectivement 5%, 6% et 3% de comprimés défectueux.
Quelle est la probabilité qu'un comprimé pris au hasard, soit défectueux ?

Solution

Posons les événements suivants :
A : "le comprimé provient de la machine A" ;
B : "le comprimé provient de la machine B" ;
C : "le comprimé provient de la machine C" ;
D : "le comprimé est défectueux".

On a

$$p(A) = 0.4,\ p(B) = 0.35,\ p(C) = 0.25,$$

$$p(D/A) = 0.05,\ p(D/B) = 0.06,\ p(D/C) = 0.03.$$

A, B et C forment un système complet d'événements, alors la probabilité qu'un comprimé pris au hasard soit défectueux (en utilisant la formule des probabilités totales) est :

$$p(D) = p(D/A)p(A) + p(D/B)p(B) + p(D/C)p(C)$$

$$p(D) == (0.05 \times 0.4) + (0.06 \times 0.35) + (0.03 \times 0.25) = 0.0485.$$

On peut construire le diagramme en arbre (l'arborescence) des événements, avec l'événement $N = \overline{D}$: "le comprimé n'est défectueux" (voir figure 3.1).

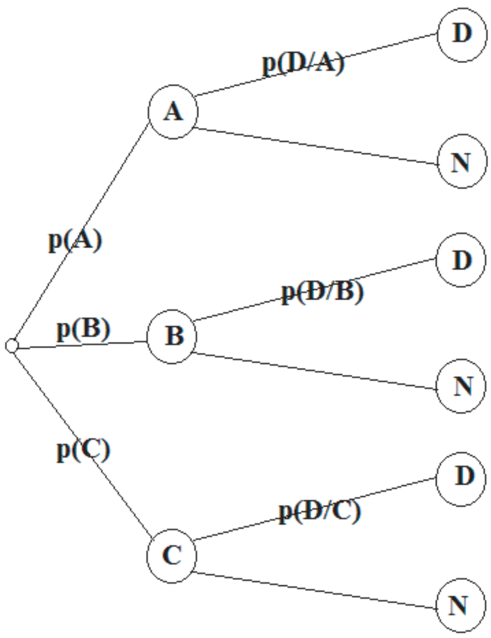


Figure 3.1 – Arborescence des événements.

3.9 Théorème de Bayes

Dans la formule des probabilités totales, on s'intéresse à la probabilité de réalisation d'un événement quelconque B, qui est donnée par

$$p(B) = p(B/A_1)p(A_1) + p(B/A_2)p(A_2) + ... + p(B/A_n)p(A_n). \tag{3.1}$$

Par contre, pour i donné, la probabilité conditionnelle de A_i sachant que l'événement B est réalisé est définie par

$$p(A_i/B) = \frac{p(A_i \cap B)}{p(B)}.$$

En utilisant la formule (3.1) et en remplaçant $p(A_i \cap B)$ par $p(A_i \cap B) = p(B/A_i)p(A_i)$, on aura le théorème suivant :

Théorème 3.2. ***Théorème de Bayes***
Soient $A_1, A_2, ..., A_n$ un système complet d'événements et B un événement quelconque. Pour tout i, $i \in \{1, 2, ..., n$, on a :

$$p(A_i/B) = \frac{p(B/A_i)p(A_i))}{p(B/A_1)p(A_1) + p(B/A_2)p(A_2) + ... + p(B/A_n)p(A_n)}.$$

Exemple 3.15. Dans l'exemple des comprimés défectueux on prend un comprimé défectueux. Quelle est la probabilité que ce défectueux provient de la machine A ?

Solution

$$p(A/D) = \frac{p(D/A)p(A)}{p(D)} = \frac{0.05 \times 0.4}{0.0485} = 0.41.$$

Chapitre 4

Variables aléatoires

Après avoir réalisé une expérience, on s'intéresse souvent à une certaine fonction du résultat et non au résultat en lui-même. Un nombre est associé à chaque résultat de l'expérience : nombre de particules émises par un élément radioactif durant un intervalle de temps donné, puissance moyenne d'un "bruit" accompagnant la réception d'un signal radio, nombre d'enfants dans une famille, etc. Ces grandeurs (ou fonctions) auxquelles on s'intéresse sont en fait des fonctions réelles définies sur l'ensemble fondamental et sont appelées **variables aléatoires**.
On considère un ensemble Ω muni d'une probabilité P.

Définition 4.1. Une variable aléatoire X est une application de l'ensemble fondamental Ω dans $\mathbb{R}$, $X : \Omega \to \mathbb{R}$, telle que que l'inverse de chaque intervalle de $\mathbb{R}$ est un événement de Ω.

On distingue deux types de variables aléatoires :

1. les variables aléatoires discrètes ;
2. les variables aléatoires continues.

4.1 Variables aléatoires discrètes

Définition 4.2. Une variable aléatoire est dite discrète si elle peut prendre un nombre fini de valeurs isolées (exemple : valeurs entières).

Exemple 4.1. En lançant un dé à six faces numérotées et en observant la face supérieure, l'ensemble fini des valeurs obtenues est : 1, 2, 3, 4, 5 et 6.

4.1.1 Loi de probabilité d'une variable aléatoire discrète

Soit X une variable aléatoire sur un ensemble fondamental Ω à valeurs finies, c'est à dire $X(\Omega) = \{x_1, x_2, ..., x_n\}$. Si l'on définit la probabilité $p(X = x_i) = p_i$ des valeurs x_i. Cette probabilité $p(X = x_i) = p_i$, est appelée la distribution ou la loi de probabilité de X, que l'on donne habituellement sous la forme du tableau suivant (tableau 4.1) :

x_i	x_1	x_2	x_3	...	x_n
$p(X = x_i)$	$p(X = x_1)$	$p(X = x_2)$	$p(X = x_3)$	...	$p(X = x_n)$

TABLE 4.1 – La loi d'une variable aléatoire

La loi de probabilité satisfait les conditions :

$$0 \leq p(X = x_i) \leq 1 \text{ et } \sum_{i=1}^{n} p(X = x_i) = 1.$$

Exemple 4.2. On jette une paire de dès bien équilibrés et on obtient l'ensemble fondamental Ω dont les éléments sont les 36 couples ordonnés des nombres allant de 1 à 6.

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

On suppose que la v.a. X est le maximum de point (a, b) de Ω, c'est-à-dire $X(a, b) = max(a, b)$; alors X sera définie par :

$$X(\Omega) = \{1, 2, 3, 4, 5, 6\}.$$

$$\begin{aligned}
p(X = 1) &= p(\{(1,1)\}) = \frac{1}{36};\\
p(X = 2) &= p(\{(1,2), (2,1), (2,2)\}) = \frac{3}{36};\\
p(X = 3) &= p(\{(1,3), (2,3), (3,1), (3,2), (3,3)\}) = \frac{5}{36};\\
p(X = 4) &= p(\{(1,4), (2,4), (3,4), (4,1), (4,2), (4,3), (4,4)\}) = \frac{7}{36};
\end{aligned}$$

de la même façon :

$$p(X = 5) = \frac{9}{36} \text{ et } p(X = 6) = \frac{11}{36}.$$

Cette information se résume dans le tableau 4.2.

On suppose maintenant une autre variable aléatoire Y, c'est la somme de composantes des couples (a, b), c'est-à-dire $Y(a, b) = a + b$; alors Y est définie par :

$$Y(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}.$$

La distribution de Y est donnée dans le tableau 4.3.

x_i	1	2	3	4	5	6
$p(X = x_i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

TABLE 4.2 – La distribution de la v.a. X.

y_i	2	3	4	5	6	7	8	9	10	11	12
$p(Y = y_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

TABLE 4.3 – La distribution de la v.a. Y.

4.1.2 Fonction de distribution et de répartition

1. **Fonction de distribution**

   Cette fonction indique la loi de probabilité de la v.a. X. Elle est représentée par un diagramme en bâtons.

   Exemple 4.3. Les diagrammes qui suivent, donnent une description graphique des distributions des variables aléatoires X (voir figure 4.1) et Y (voir figure 4.2) de l'exemple précédent.

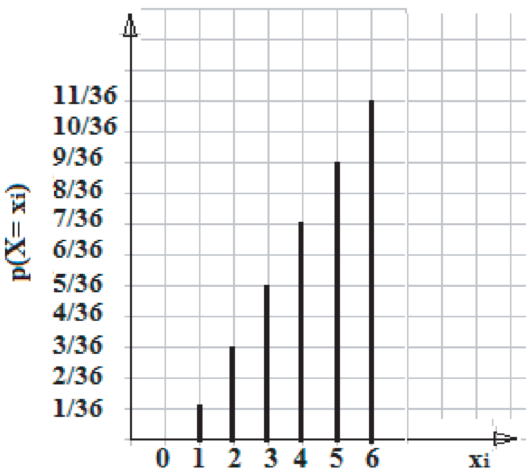


FIGURE 4.1 – Distribution de la v.a. X.

2. **Fonction de répartition**

   La fonction de répartition donne la probabilité que la variable aléatoire X prenne une valeur inférieure à x. La fonction de répartition est définie par :

$$F(x) = p(X \leq x) = \sum_{x_i \leq x} p(X = x_i).$$

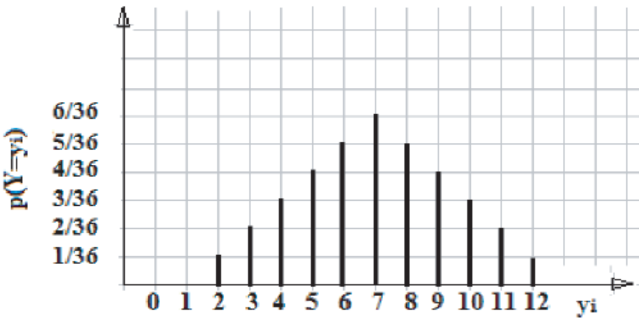


FIGURE 4.2 – Distribution de la v.a. Y.

Remarque 4.1. La représentation graphique de la fonction de répartition de le cas discret prend la forme d'un diagramme en escaliers (voir figure 4.3).
F est monotone croissante et prend ses valeurs dans $[0, 1]$.

Exemple 4.4. La fonction de répartition de la variable aléatoire X est donnée par :

$$F(x) = \begin{cases} 0, & \text{si } x < 1\,; \\ \frac{1}{36}, & \text{si } 1 \leq x < 2\,; \\ \frac{4}{36}, & \text{si } 2 \leq x < 3\,; \\ \frac{9}{36}, & \text{si } 3 \leq x < 4\,; \\ \frac{16}{36}, & \text{si } 4 \leq x < 5\,; \\ \frac{25}{36}, & \text{si } 5 \leq x < 6\,; \\ \frac{36}{36} = 1, & \text{si } x \geq 6. \end{cases}$$

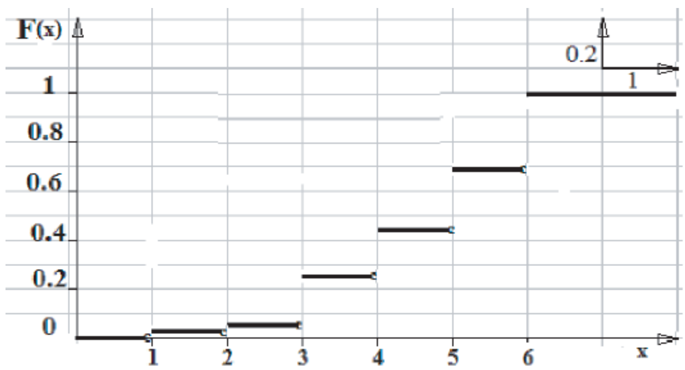


FIGURE 4.3 – Courbe de la fonction de répartition de la v.a. X.

4.1.3 Espérance mathématique d'une variable aléatoire discrète

Définition 4.3. Soit X une variable aléatoire ayant la loi de probabilité $p(X = x_i)_{i=\overline{1,n}}$. La moyenne ou l'espérance mathématique de X que l'on note $E(X)$ ou μ_x est la somme des valeurs prises par X pondérées par les probabilités qui leur sont associées (la valeur prise en moyenne par cette v.a.), elle est donnée par :

$$\begin{aligned} E(X) &= x_1p(X=x_1)+x_2p(X=x_2)+...+x_np(X=x_n) \\ &= \sum_{i=1}^{n} x_ip(X=x_i). \end{aligned}$$

Exemple 4.5. On reprend l'exemple 4.2.
• L'espérance mathématique de la variable aléatoire X est :

$$\begin{aligned} E(X) &= \sum_{i=1}^{n} x_ip(X=x_i) \\ &= 1.\frac{1}{36}+2.\frac{3}{36}+3.\frac{5}{36}+4.\frac{7}{36}+5.\frac{9}{36}+6.\frac{11}{36} \\ &= \frac{161}{36}=4,47 \end{aligned}$$

• L'espérance mathématique de la variable aléatoire Y est :

$$\begin{aligned} E(Y) &= \sum_{i=1}^{n} y_ip(Y=y_i) \\ &= 2.\frac{1}{36}+3.\frac{2}{36}+4.\frac{3}{36}+5.\frac{4}{36}+6.\frac{5}{36}+7.\frac{6}{36}+8.\frac{5}{36}+9.\frac{4}{36}+10.\frac{3}{36}+11.\frac{2}{36}+12.\frac{1}{36} \\ &= \frac{252}{36}=7. \end{aligned}$$

Propriétés de $E(X)$

Désignons par X et Y deux variables aléatoires définies sur Ω, α et β deux constantes réelles.

1. $E(\alpha X)=\alpha E(X)$, $\forall\alpha\in\mathbb{R}$.
2. $E(\alpha X+\beta)=\alpha E(X)+\beta$, $\forall\alpha,\beta\in\mathbb{R}$.
3. $E(X+Y)=E(X)+E(Y)$.
4. $E(\alpha X+\beta Y)=\alpha E(X)+\beta E(Y)$, $\forall\alpha,\beta\in\mathbb{R}$.

4.1.4 Variance et écart type d'une variable aléatoire discrète

Définition 4.4. La moyenne d'une variable aléatoire X mesure, dans un certain sens, la valeur moyenne de X et la variance (ou sa racine carrée est l'écart-type) exprime à quel point les valeurs prises par une variable aléatoire X sont dispersées autour de la moyenne.

1. **Variance de la variable aléatoire X**
   La variance de X, que l'on note $V(X)$ est définie par
   $$V(X) = E[(X - E(X))^2] = \sum_{i=1}^{n} (x_i - E(X))^2 p(X = x_i)$$
   ou
   $$V(X) = E(X^2) - (E(X))^2 = \sum_{i=1}^{n} x_i^2 p(X = x_i) - (\sum_{i=1}^{n} x_i p(X = x_i))^2.$$
2. **Écart type de la variable aléatoire X**
   L'écart type de X, que l'on note $\sigma(X)$ est la racine carrée de $V(X)$:
   $$\sigma(X) = \sqrt{V(X)}.$$

Exemple 4.6. Considérons les variables aléatoires X et Y de l'exemple 4.2, avec leurs moyennes $E(X) = 4,47$ et $E(Y) = 7$.

- La variance de la variable aléatoire X est donnée par $V(X) = E(X^2) - (E(X))^2$, avec

$$\begin{aligned} E(X^2) &= \sum_{i=1}^{n} x_i^2 p(X = x_i) \\ &= 1^2.\frac{1}{36} + 2^2.\frac{3}{36} + 3^2.\frac{5}{36} + 4^2.\frac{7}{36} + 5^2.\frac{9}{36} + 6^2.\frac{11}{36} \\ &= \frac{191}{36} = 21,97. \end{aligned}$$

Par conséquent $V(X) = E(X^2) - (E(X))^2 = 21,97 - (4,47)^2 = 1,99$ et $\sigma(X) = \sqrt{1,99} = 1,4$.

- La variance de la variable aléatoire Y est donnée par $V(Y) = E(Y^2) - (E(Y))^2$, avec

$$\begin{aligned} E(Y^2) &= \sum_{i=1}^{n} y_i^2 p(Y = y_i) \\ &= 2^2\frac{1}{36} + 3^2\frac{2}{36} + 4^2\frac{3}{36} + 5^2\frac{4}{36} + 6^2\frac{5}{36} + 7^2\frac{6}{36} + 8^2\frac{5}{36} + 9^2\frac{4}{36} + 10^2\frac{3}{36} + 11^2\frac{2}{36} + 12^2\frac{1}{36} \\ &= = 54,8. \end{aligned}$$

Par conséquent $V(Y) = E(Y^2) - (E(Y))^2 = 54,8 - (7)^2 = 5,8$ et $\sigma(Y) = \sqrt{5,8} = 2,4$.

Propriétés de $V(X)$ et $\sigma(X)$

Désignons par X une variable aléatoire définie sur Ω, α et β deux constantes réelles.

1. La variance d'une constante est nulle : $V(\alpha) = 0$, $\forall \alpha \in \mathbb{R}$.
2. $V(X + \alpha) = V(X)$, $\forall \alpha \in \mathbb{R}$.
3. $V(\alpha X) = \alpha^2 V(X)$, $\forall \alpha \in \mathbb{R}$.
4. $V(\alpha X + \beta) = \alpha^2 V(X)$, $\forall \alpha, \beta \in \mathbb{R}$.

D'où

1. $\sigma(X + \alpha) = \sigma(X)$, $\forall \alpha \in \mathbb{R}$.
2. $\sigma(\alpha X) = |\alpha| \sigma(X)$, $\forall \alpha \in \mathbb{R}$.

Remarque 4.2. Soit X une variable aléatoire de moyenne μ et d'écart type σ, on définit la variable aléatoire centrée réduite X^* correspondant à X par

$$X^* = \frac{X - \mu}{\sigma},$$

avec $E(X^*) = 0$ et $V(X^*) = 1$.

4.2 Variables aléatoires continues

Une variable aléatoire est dite continue si elle peut prendre toutes valeurs comprises dans un intervalle $]a, b]$.

Exemple 4.7. Le poids d'un enfant à la naissance est compris entre $2,7$ kg et $5,6$ kg.

4.2.1 Fonction de répartition

La fonction de répartition d'une variable continue X est définie par

$$F_X(x) = p(X \leq x).$$

La fonction F_X indique la probabilité que X soit strictement inférieure à tout x de l'intervalle de définition.

Propriétés

1. $F_X(x)$ est positive et : $\lim\limits_{x \to -\infty} F_X(x) = 0$, $\lim\limits_{x \to +\infty} F_X(x) = 1$.
2. Si la fonction F_X est continue et admet une dérivée, la variable aléatoire est dite absolument continue.

3. La représentation graphique de F_X prend la forme d'une courbe cumulative.

Remarque 4.3. Dans le cas d'**une variable aléatoire continue**, on a :

1. La probabilité attachée à un point x est nulle : $p(X = x) = 0$.
2. $p(X \leq x) = p(X < x) + p(X = x) = p(X < x)$.
3. La probabilité que la v.a. $X \in [a, b]$ est donnée par :

$$\begin{aligned} p(a \leq X \leq b) &= p(a < X \leq b) \\ &= p(a \leq X < b) \\ &= p(a < X < b) \\ &= p(X < b) - p(X < a) \\ &= F(b) - F(a). \end{aligned}$$

4.2.2 Densité de probabilité

Soit X une variable aléatoire dont l'ensemble de valeurs $X(\Omega)$ est l'intervalle $[a, b]$. Rappelons que par définition

$$p(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a).$$

La fonction f est la distribution (densité de probabilité) de la variable aléatoire continue X. Cette fonction satisfait les conditions suivantes :

1. $f(x) \geq 0$ et $F(x) = \int_{-\infty}^{x} f(t)dt$.
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$.

4.2.3 Espérance mathématique et variance d'une variable aléatoire continue

Soit X une variable aléatoire de densité de probabilité f, dont le domaine de définition est $]-\infty, +\infty[$.

1. **Espérance mathématique**
   L'espérance mathématique de la v.a. X est définie par :

$$E(X) = \int_{-\infty}^{+\infty} x f(x)dx$$

2. **Variance et écart type**
   La variance de la v.a. X est définie par :

$$\begin{aligned} V(X) &= \int_{-\infty}^{+\infty} (x - E(X))^2 f(x)dx \\ &= \int_{-\infty}^{+\infty} x^2 f(x)dx - (E(X))^2 \\ &= E(X^2) - (E(X))^2. \end{aligned}$$

   Par définition, l'écart type est donné par $\sigma(X) = \sqrt{V(X)}$.

Exemple 4.8. Soit X une variable aléatoire ayant une densité de probabilité (fonction de distribution) :

$$f(x) = \begin{cases} \frac{1}{2}(2 - x), & \text{si } 0 \leq x \leq 2\,; \\ 0, & \text{si } x \in]-\infty, 0[\cup]2, +\infty[. \end{cases}$$

1. La densité de probabilité vérifie :

$$f(x) \geq 0, \forall x \in [0, 2] \text{ et } \int_{-\infty}^{+\infty} f(x)dx = 1.$$

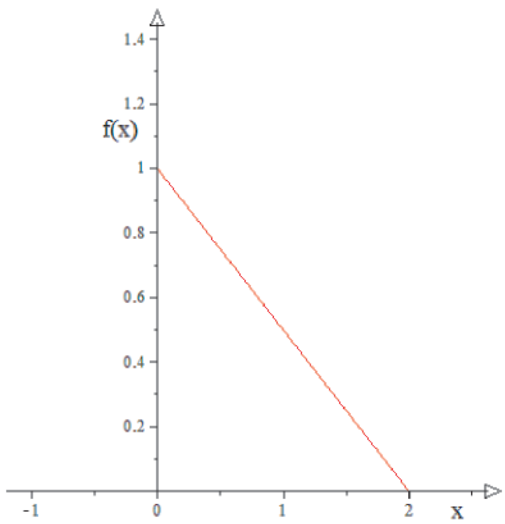


FIGURE 4.4 – Densité de probabilité de la v.a. X.

En effet,

$\forall x \in [0,2]$, on a $0 \leq f(x) \leq 1$ et $f(x) = 0$ ailleurs ;

$$\begin{aligned}
\int_{-\infty}^{+\infty} f(x)dx &= \int_0^2 f(x)dx \\
&= \int_0^2 \frac{1}{2}(2-x)dx \\
&= \frac{1}{2}\int_0^2 (2-x)dx \\
&= \frac{1}{2}[2x - \frac{x^2}{2}]_0^2 \\
&= \frac{1}{2}(4-2-0) = 1.
\end{aligned}$$

2. La fonction de répartition F est :

$$F_X(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, & \text{si } x < 0\,; \\ x - \frac{1}{4}x^2, & \text{si } 0 \leq x \leq 2\,; \\ 1, & \text{si } x > 2. \end{cases}$$

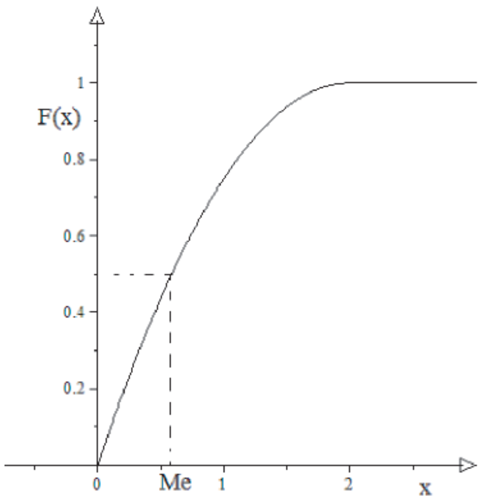


FIGURE 4.5 – Fonction de répartition de la v.a. X.

3. L'espérance de X :

$$\begin{aligned} E(X) = \int_{-\infty}^{+\infty} xf(x)dx &= \int_0^2 xf(x)dx \\ &= \int_0^2 \frac{1}{2}x(2-x)dx \\ &= \frac{1}{2}\int_0^2 (2x - x^2)dx \\ &= \frac{1}{2}[x^2 - \frac{x^3}{3}]_0^2 \\ &= \frac{2}{3} - 0 = \frac{2}{3}. \end{aligned}$$

4. La variance de X :

$$\begin{aligned} V(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - E(X)^2 &= \int_0^2 x^2 f(x)dx - E(X)^2 \\ &= \int_0^2 \frac{1}{2}x^2(2-x)dx - (\frac{2}{3})^2 \\ &= \frac{1}{2}\int_0^2 (2x^2 - x^3)dx - (\frac{2}{3})^2 \\ &= \frac{1}{2}[\frac{2}{3}x^3 - \frac{x^4}{4}]_0^2 - (\frac{2}{3})^2 \\ &= \frac{2}{3} - \frac{4}{9} = \frac{2}{9}. \end{aligned}$$

5. L'écart type de X : $\sigma(X) = \sqrt{V(X)} = \sqrt{\frac{2}{9}} = \frac{\sqrt{2}}{3}$.

4.2.4 Médiane et mode d'une variable aléatoire continue

La médiane

La médiane d'une variable aléatoire continue est le nombre réel Me tel que

$$F(Me) = 0.5 = \frac{1}{2}.$$

Autrement dit, la médiane c'est la valeur Me de X tel que $p(X < Me) = 0.5$.

Le mode

Le mode Mo est la valeur de X, qui correspond à un maximum de la fonction de densité. Il peut exister plusieurs modes, si il existe un seul maximum la densité est dite unimodale.

Exemple 4.9. Dans l'exemple 4.8, on a

$$F(x) = x - \frac{1}{4}x^2 = \frac{1}{2} \Rightarrow x = \frac{4-\sqrt{8}}{2} \text{ ou } x = \frac{4+\sqrt{8}}{2}.$$

$$x = \frac{4-\sqrt{8}}{2} \simeq 0,5857864376 \in [0;2]$$

et

$$x = \frac{4+\sqrt{8}}{2} \simeq 3.414213562 \notin [0;2],$$

alors

$$Me = \frac{4-\sqrt{8}}{2}.$$

f admet un maximum pour la valeur de $x = 0$, alors $Mo = 0$.

Chapitre 5

Lois usuelles de probabilités

5.1 Lois de probabilités discrètes

5.1.1 Loi de Bernoulli

Une expérience aléatoire ayant deux résultats possibles (succès et échec) est appelée experience de Bernoulli.
Si A est l'événement succès et $\overline{A}$ est l'événement échec, on a :

$$\begin{aligned} p(A) &= p,\ 0 \leq p \leq 1; \\ p(\overline{A}) &= 1 - p = q,\ 0 \leq q \leq 1. \end{aligned}$$

On dit que X suit une loi de Bernoulli de probabilité p, et on note

$$X \rightsquigarrow \text{Bernoulli}(p)$$

Espérance mathématique

La variable aléatoire X prend deux valeurs possibles $\{0; 1\}$: 0 en cas d'échec et 1 en cas de réussite et son espérance mathématique est :

$$E(X) = p.$$

En effet,

$$E(X) = \sum_{i=0}^{1} x_i p(X = x_i) = 0.q + 1.p = p.$$

Variance

La variance de cette variable aléatoire qui suit une loi de Bernoulli(p) est :

$$V(X) = p.q$$

En effet,

$$\begin{aligned}
V(X) &= E(X^2) - (E(X)^2) \\
&= \sum_{i=0}^{1} x_i^2 p(X = x_i) - (E(X)^2) \\
&= (0^2 q + 1^2 p) - p^2 \\
&= p - p^2 \\
&= p(1-p) = p.q
\end{aligned}$$

Écart type

L'écart type est

$$\sigma(X) = \sqrt{V(X)} = \sqrt{p.q}.$$

La loi de Bernoulli(p) se résume dans le tableau 5.1 :

k	0	1		$E(X) = p$
$p(X = k)$	$p(X = 0) = q$	$p(X = 1) = p$		$V(X) = p.q$
	$0 \leq q \leq 1$	$0 \leq p \leq 1$		$\sigma(X) = \sqrt{p.q}$

TABLE 5.1 – Loi de Bernoulli(p).

Exemple 5.1. On jette un dé équilibré et on s'intéressera au résultat "avoir le chiffre 2". A : "obtenir le chiffre 2 sur la surface supérieure du dé".

$$\begin{aligned}
A = \{2\} \quad &\text{et} \quad \overline{A} = \{1, 3, 4, 5, 6\}; \\
p(A) = p = \frac{1}{6} \quad &\text{et} \quad p(\overline{A}) = 1 - p = q = \frac{5}{6}.
\end{aligned}$$

$$\begin{aligned}
\text{Le jet d'un dé} \quad \text{est} \quad & \text{une expérience de Bernoulli;} \\
\text{avec} \quad & p = \frac{1}{6} \text{ et } q = \frac{5}{6}; \\
& E(X) = p = \frac{1}{6}; \\
& V(X) = p.q = \frac{1}{6} \times \frac{5}{6} = \frac{5}{36}; \\
& \sigma(X) = \sqrt{V(X)} = \sqrt{p.q} = \frac{\sqrt{5}}{6}.
\end{aligned}$$

5.1.2 Loi binomiale

Soit une expérience de Bernoulli répétée n fois dans les mêmes conditions et de manières indépendantes (X_1, X_2, ... , X_n). La loi binomiale, notée $\mathcal{B}(n,p)$, $X = \sum\limits_{i=1}^{n} X_i = X_1 + X_2 + ... + X_n$ modélise le nombre de succès obtenus lors de la répétition indépendante de plusieurs expériences aléatoires identiques (avec p la probabilité du succès et $q = 1-p$ la probabilité de l'échec).
La variable aléatoire X correspond au nombre de succès, si on a k succès on aura $(n-k)$ échecs et la probabilité d'avoir k succès dans une experience aléatoire répétée n fois, est :

$$p(X = k) = C_n^k.p^k.q^{(n-k)},$$

avec $C_n^k = \frac{n!}{(n-k)!k!}$.
Il est facile de démontrer que l'on a bien une loi de probabilité, car :

$$\sum_{k=0}^{n} p(X = k) = \sum_{k=0}^{n} C_n^k p^k q^{(n-k)} = (p+q)^n = 1, \text{ car } p + q = 1.$$

Remarque 5.1. Le développement du binôme de Newton $(p+q)^n$ permet d'obtenir l'ensemble des probabilités pour une distribution binomiale $\mathcal{B}(n,p)$ avec n et p des valeurs données.

Espérance mathématique

L'espérance mathématique de la distribution binomiale $\mathcal{B}(n,p)$ est :

$$E(X) = n.p$$

En effet,

$$\begin{aligned} E(X) &= E(\sum_{i=1}^{n} X_i) \text{ où chaque } X_i \text{ est une v.a. de Bernoulli} \\ &= E(X_1 + X_2 + ... + X_n) \\ &= E(X_1) + E(X_2) + ... + E(X_n) \\ &= \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p = np. \end{aligned}$$

Variance

La variance de cette variable aléatoire qui suit une loi de binomiale $\mathcal{B}(n,p)$ est :

$$V(X) = n.p.q$$

En effet,

$$\begin{aligned} V(X) &= V(\sum_{i=1}^{n} X_i) \text{ où chaque } X_i \text{ est une v.a. de Bernoulli} \\ &= V(X_1 + X_2 + ... + X_n) \\ &= V(X_1) + V(X_2) + ... + V(X_n) \\ &= \sum_{i=1}^{n} V(X_i) = \sum_{i=1}^{n} pq = npq. \end{aligned}$$

Écart type

L'écart type est

$$\sigma(X) = \sqrt{V(X)} = \sqrt{n.p.q}.$$

La loi binomiale $\mathcal{B}(n,p)$ se résume dans le tableau 5.2 :

Réalisations	$k = 0, 1, 2, ..., n$		$E(X) = np$
Probabilité d'avoir k réussites	$p(X = k) = C_n^k.p^k.q^{(n-k)}$		$V(X) = npq$
avec	$0 \leq p \leq 1$ et $0 \leq q \leq 1$		$\sigma(X) = \sqrt{npq}$

TABLE 5.2 – Loi binomiale $\mathcal{B}(n,p)$.

Exemple 5.2. On reprend l'exemple du dé et on refait l'expérience du jet 5 fois (on jette le dé 5 fois). On s'intéressera au nombre de fois, où on obtient un 2.
A : "obtenir un 2 sur la surface supérieure du dé".
$p(A) = p = \frac{1}{6}$ et $p(\overline{A}) = q = 1 - p = \frac{5}{6}$.
X suit une loi Binomiale $\mathcal{B}(n,p) = \beta(5, \frac{1}{6})$.

$$p(X = k) = C_n^k.p^k.q^{(n-k)} = C_5^k.(\frac{1}{6})^k.(\frac{5}{6})^{(5-k)}.$$

1. Quelle est la probabilité d'obtenir :
   (a) deux fois le chiffre 2 ?
   (b) au moins trois fois le chiffre 2 ?
2. Calculer $E(X)$, $V(X)$ et $\sigma(X)$.

Solution :

1. La probabilité d'obtenir :

(a) deux fois le chiffre 2 est :

$$p(X=2)=C_5^2(\frac{1}{6})^2(\frac{5}{6})^{(5-2)}=10\times 0,028\times 0,58=0,16.$$

(b) au moins trois fois le chiffre 2 est :

$$\begin{aligned} p(X\geq 3) &= p(X=3)+p(X=4)+p(X=5)\\ &= 1-p(X<3)\\ &= 1-[p(X=0)+p(X=1)+p(X=2)]\\ &= 1-[C_5^0(\frac{1}{6})^0(\frac{5}{6})^{(5-0)}+C_5^1(\frac{1}{6})^1(\frac{5}{6})^{(5-1)}+C_5^2(\frac{1}{6})^2(\frac{5}{6})^{(5-2)}]\\ &= 1-[0,4+0,4+0,16]=0,036. \end{aligned}$$

2. Calcul de $E(X)$, $V(X)$ et $\sigma(X)$:

$$\begin{aligned} E(X) &= np=5\times\frac{1}{6}=\frac{5}{6}.\\ V(X) &= npq=5\times\frac{1}{6}\times\frac{5}{6}=\frac{25}{36}.\\ \sigma(X) &= \sqrt{V(X)}=\frac{5}{6}. \end{aligned}$$

Théorème 5.1. *Si $X \rightsquigarrow \mathcal{B}(n,p)$ et $Y \rightsquigarrow \mathcal{B}(m,p)$ sont deux variables aléatoires indépendantes de même probabilité p, alors leurs somme $X+Y$ est une variable aléatoire qui suit une loi binomiale :*

$$X+Y\rightsquigarrow\mathcal{B}(n+m,p).$$

5.1.3 Loi de Poisson

On dit qu'une variable aléatoire X suit une loi de poisson (appelée aussi loi des événements rares ou de petits nombres) de paramètre réel λ, notée $\mathcal{P}(\lambda)$, si elle prend des valeurs entières dont les probabilités de réalisation sont :

$$\forall k\in\mathbb{N},\ p(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\ e=2,718...$$

Paramètres de la loi de Poisson $\mathcal{P}(\lambda)$:

Espérance mathématique : $E(X)=\lambda$.
Variance : $V(X)=\lambda$.
Écart type : $\sigma(X)=\sqrt{\lambda}$.

Exemple 5.3. Une centrale téléphonique reçoit en moyenne 300 appels par heure. Quelle est la probabilité que durant une minute, la centrale reçoit exactement deux appels ?
Solution :
Les appels dans cette centrale suivent une loi de poisson de paramètre $\lambda = \frac{300}{60} = 5$ appels par minutes en moyenne.

$$\begin{aligned} p(X=2) &= e^{-\lambda}\frac{\lambda^k}{k!} = e^{-5}\frac{5^2}{2!} \\ &= 0,08422. \end{aligned}$$

5.1.4 Table de la loi de Poisson

A l'intersection de la colonne λ et de la ligne k, figure la probabilité pour que la variable de Poisson Y de paramètre λ soit égale à la valeur entière k : $p(Y=k) = e^{-\lambda}.\frac{\lambda^k}{k!}$

Probabilités individuelles $P_\lambda(k) = P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}$

k \ λ	0,02	0,04	0,06	0,08	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5	1	1,5	2	2,5	3	4	5
0	0,980	0,961	0,942	0,923	0,905	0,861	0,819	0,779	0,741	0,705	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
1	0,020	0,038	0,057	0,074	0,090	0,129	0,164	0,195	0,222	0,247	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
2	0,000	0,001	0,002	0,003	0,005	0,010	0,016	0,024	0,033	0,043	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
3	0,000	0,000	0,000	0,000	0,000	0,000	0,001	0,002	0,003	0,005	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
9	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
10	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
11	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
12	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
13	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
14	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

FIGURE 5.1 – Table de la loi de Poisson.

Exemple 5.4. Sur une autoroute, il y a en moyenne deux accidents par semaine.
Quelle est la probabilité qu'il y aura cinq accidents durant un week-end ?
Solution : La loi de X du nombre d'accidents sur cette route suit une loi de Poisson de paramètre $\lambda = 2$ et la probabilité qu'il y aura cinq accidents durant un week-end est

$$p(X=5) = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-2}\frac{2^5}{5!} = 0,0361.$$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
0	0,368	0,135	0,050	0,018																
1	0,368	0,271	0,149	0,073	0,034	0,015														
2	0,184	0,271	0,224	0,147	0,084	0,045	0,022	0,011												
3	0,061	0,180	0,224	0,195	0,140	0,089	0,052	0,029	0,015											
4	0,015	0,090	0,168	0,195	0,175	0,134	0,091	0,057	0,034	0,019	0,010									
5		0,036	0,101	0,156	0,175	0,161	0,128	0,092	0,061	0,038	0,022	0,013								
6		0,012	0,050	0,104	0,146	0,161	0,149	0,122	0,091	0,063	0,041	0,025	0,015							
7			0,022	0,060	0,104	0,138	0,149	0,140	0,117	0,090	0,065	0,044	0,028	0,017	0,010					
8				0,030	0,065	0,103	0,130	0,140	0,132	0,113	0,089	0,066	0,046	0,030	0,019	0,012				
9				0,013	0,036	0,069	0,101	0,124	0,132	0,125	0,109	0,087	0,066	0,047	0,032	0,021	0,014			
10					0,018	0,041	0,071	0,099	0,119	0,125	0,119	0,105	0,086	0,066	0,049	0,034	0,023	0,015		
11						0,023	0,045	0,072	0,097	0,114	0,119	0,114	0,101	0,084	0,066	0,050	0,036	0,025	0,016	0,011
12						0,011	0,026	0,048	0,073	0,095	0,109	0,114	0,110	0,098	0,083	0,066	0,050	0,037	0,026	0,018
13							0,014	0,030	0,050	0,073	0,093	0,106	0,110	0,106	0,096	0,081	0,066	0,051	0,038	0,027
14								0,017	0,032	0,052	0,073	0,090	0,102	0,106	0,102	0,093	0,080	0,065	0,051	0,039
15									0,019	0,035	0,053	0,072	0,088	0,099	0,102	0,099	0,091	0,079	0,065	0,052
16									0,011	0,022	0,037	0,054	0,072	0,087	0,096	0,099	0,096	0,088	0,077	0,065
17										0,013	0,024	0,038	0,055	0,071	0,085	0,093	0,096	0,094	0,086	0,076
18											0,015	0,026	0,040	0,055	0,071	0,083	0,091	0,094	0,091	0,084
19												0,016	0,027	0,041	0,056	0,070	0,081	0,089	0,091	0,089
20													0,018	0,029	0,042	0,056	0,069	0,080	0,087	0,089
21													0,011	0,019	0,030	0,043	0,056	0,068	0,078	0,085
22														0,012	0,020	0,031	0,043	0,056	0,068	0,077
23															0,013	0,022	0,032	0,044	0,056	0,067
24																0,014	0,023	0,033	0,044	0,056
25																	0,015	0,024	0,034	0,045
26																	0,010	0,016	0,025	0,034
27																		0,011	0,017	0,025
28																			0,012	0,018
29																				0,013
30																				

FIGURE 5.2 – Table de la loi de Poisson pour k variant de 0 à 30.

5.1.5 L'approximation de la loi binomiale par la loi de Poisson

Considérons une variable aléatoire X suivant une loi binomiale $\mathcal{B}(n;p)$. Si n tend vers l'infini et p tend vers 0 ($n \to +\infty$ et $p \to 0$), la loi binomiale converge vers une loi de Poisson.
En pratique, si $n > 25$ et $np < 5$, alors la loi binomiale $\mathcal{B}(n;p)$ est approchée par la loi de poisson $\mathcal{P}(\lambda)$, avec $\lambda = np$.

Exemple 5.5. Des observations ont montré que la probabilité qu'un homme soit atteint d'une maladie M est $p = 0,1$. En considérant 40 hommes pris au hasard, soit k le nombre d'hommes touchés par la maladie et X la variable aléatoire qui compte le nombre d'hommes malades.

1. Quelle est la loi que suit la v.a. X ? Donner sa loi de distribution $p(X = k)$.
2. Calculer $E(X)$, $V(X)$ et $\sigma(X)$.
3. Par quelle loi peut-on approcher la loi de X ?
4. Calculer $p(X = 2)$ et $p(X = 5)$.

Solution :

1. X suit une loi binomiale $\mathcal{B}(40; 0.1)$:
$$p(X = k) = C_{40}^k.(\frac{1}{10})^k.(\frac{9}{10})^{(40-k)}$$
2. Calcul de $E(X)$, $V(X)$ et $\sigma(X)$:
$$\begin{aligned} E(X) &= np = 40 \times \frac{1}{10} = 4. \\ V(X) &= npq = 40 \times \frac{1}{10} \times \frac{9}{10} = 3,6. \\ \sigma(X) &= \sqrt{V(X)} = 1,89. \end{aligned}$$
3. Cette loi binomiale $\mathcal{B}(40; 0.1)$ est approchée par une loi de Poisson $\mathcal{P}(\lambda)$:

   $(n = 40 > 25$ et $np = 4 < 5) \Rightarrow$ Binomiale $\mathcal{B}(40; 0.1) \rightsquigarrow$ Poisson $\mathcal{P}(\lambda), \lambda = np = 4$.
$$p(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} = e^{-4}\frac{4^k}{k!}.$$
4. Calcul de $p(X = 2)$ et $p(X = 5)$:
$$p(X = 2) = e^{-4}\frac{4^2}{2!} = 0,14653.$$
$$p(X = 5) = e^{-4}\frac{4^5}{5!} = 0,15629.$$

5.2 Lois de probabilités continues

5.2.1 Loi Normale

En théorie des probabilités et en statistique, la loi normale est l'une des lois de probabilité les plus adaptées pour modéliser des phénomènes naturels issus de plusieurs événements aléatoires. Elle est en lien avec de nombreux objets mathématiques dont le mouvement brownien, le bruit blanc gaussien ou d'autres lois de probabilité. Elle est également appelée loi gaussienne, loi de Gauss ou loi de Laplace-Gauss des noms de Laplace (1749-1827) et Gauss (1777-1855), deux mathématiciens, astronomes et physiciens qui l'ont étudiée.

Soit X une variable aléatoire. On dit que X suit une loi normale ou Laplace-Gauss de paramètres m (ou μ) et σ ($m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$), si sa densité de probabilité est définie par :

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

avec $x \in \mathbb{R}$, $E(X) = m$ et σ est l'écart type de la v.a. X.

La courbe de cette densité de probabilité est appelée courbe de Gauss ou courbe en cloche. La figure 5.3 donne une illustration de quelque densités de probabilités en variant leurs paramètres correspondants (m et σ) et l'aire sous la courbe est toujours égale à 1.

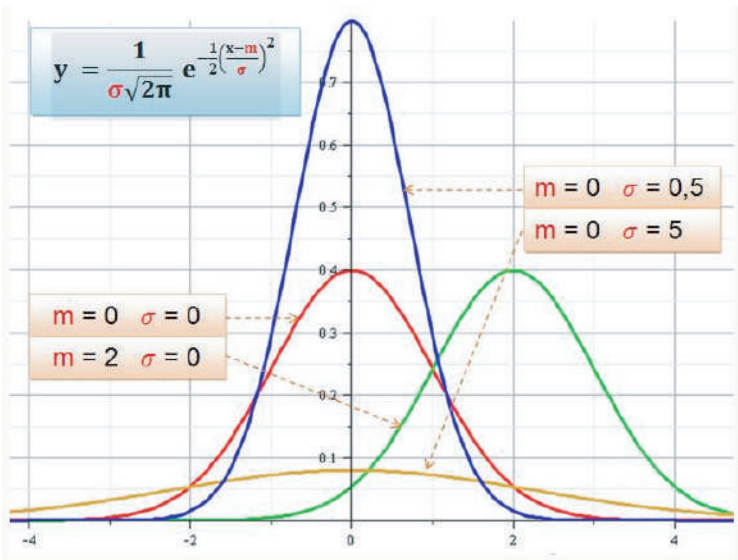
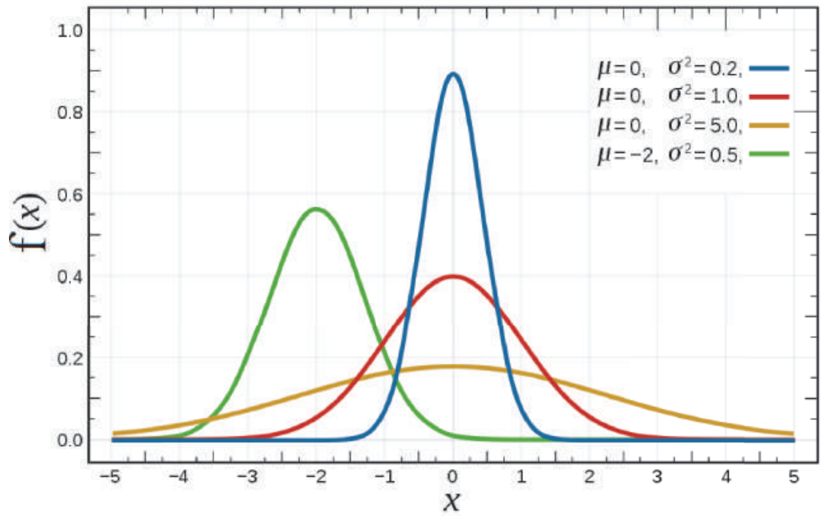


FIGURE 5.3 – Illustration de lois normales avec variations de m et σ.

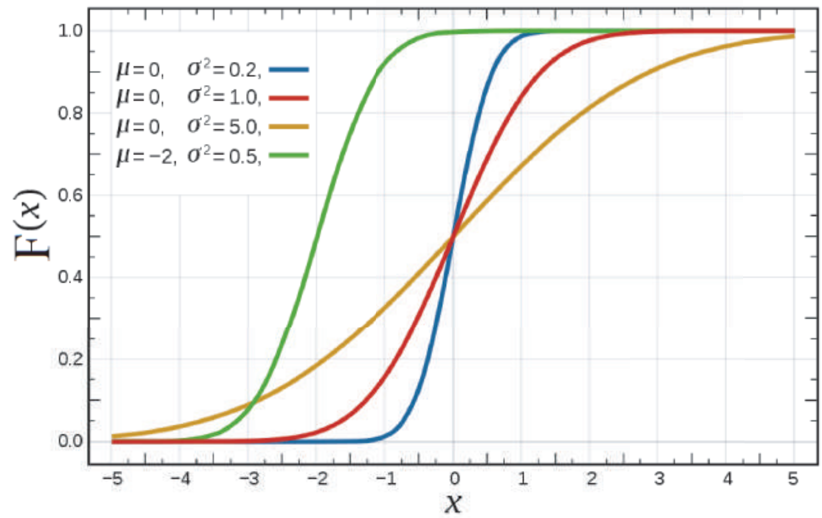
Fonction de répartition

La fonction de répartition F de la loi normale $\mathcal{N}(m, \sigma)$ est :

$$F(x) = p(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-1}{2}(\frac{t-m}{\sigma})^2} dt.$$



(a) Densités de probabilités



(b) Fonctions de répartitions

Figure 5.4 – Fonctions de répartition de la loi Normale $\mathcal{N}(\mu, \sigma)$ avec variation des paramètres μ et σ. La courbe en couleur rouge est associée à la loi Normale centrée réduite $\mathcal{N}(0, 1)$.

5.2.2 Loi Normale centrée réduite

Dans la pratique, on rencontre très souvent la loi normale. Afin d'éviter le calcul numérique de la fonction de répartition pour chaque application, on utilisera la loi normale centrée réduite, dont les valeurs existent et sont tablées.

Définition 5.1. Variable aléatoire centrée réduite

1. Une variable aléatoire centrée est une v.a. dont l'espérance est nulle $E(X) = 0$.
2. Une variable aléatoire réduite est une v.a. dont l'écart type $\sigma(X) = 1$ $(V(X) = 1)$.
3. La variable aléatoire $\frac{X-E(X)}{\sigma(X)}$ est une v.a. centrée réduite.
   En effet,

$$\begin{aligned} E(\frac{X-E(X)}{\sigma(X)}) &= \frac{1}{\sigma(X)}(E(X)-E(X)) = 0; \\ V(\frac{X-E(X)}{\sigma(X)}) &= \frac{1}{\sigma^2(X)}(V(X)) = \frac{\sigma^2(X)}{\sigma^2(X)} = 1. \end{aligned}$$

Théorème 5.2. *Soit X une variable aléatoire continue suivant une loi normale $\mathcal{N}(m,\sigma)$. Si on applique le changement de variable $Z = \frac{X-m}{\sigma}$ et le changement de bornes correspondantes $z = \frac{x-m}{\sigma}$, on a :*

$$F_X(x) = p(X \leq x) = p(\frac{X-m}{\sigma} \leq \frac{x-m}{\sigma}) = p(Z \leq z) = F_Z(z).$$

La variable aléatoire Z suit une loi normale centrée réduite $\mathcal{N}(0,1)$.

$$\text{Si } X \rightsquigarrow \mathcal{N}(m,\sigma) \text{ alors } Z = \frac{X-m}{\sigma} \rightsquigarrow \mathcal{N}(0,1).$$

Densité de probabilité de Z

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

Fonction de répartition de Z

$$F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-1}{2}t^2} dt.$$

Propriété 5.1. *Si la variable aléatoire Z suit une loi normale centrée réduite $\mathcal{N}(0,1)$, alors :*

1. $f(z) = f(-z)$ *;*

2. $\int_{-\infty}^{0} f(t)dt = \int_{0}^{+\infty} f(t)dt = \frac{1}{2}\int_{-\infty}^{+\infty} f(t)dt = \frac{1}{2}$ *;*

3. $\forall z \in \mathbb{R}_+$*, on a* $\int_{-\infty}^{-z} f(t)dt = \int_{z}^{+\infty} f(t)dt$ *;*

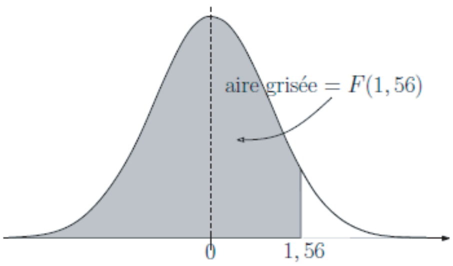
4. $\forall z \in \mathbb{R}_+$*, on a* $F_Z(-z) = 1 - F_Z(z)$*.*

Exemple 5.6. On suppose qu'une certaine variable aléatoire Z suit une loi normale centrée réduite $\mathcal{N}(0,1)$.

1. Pour quelle proportion d'individus on a $Z \leq 1,56$?
2. Pour quelle proportion d'individus on a $Z \geq 1,49$?
3. Calculer $p(Z \leq -1,1)$.

Solution

1. Calcul de $p(Z \leq 1,56)$:



La valeur de $p(Z \leq 1,56) = F(1,56)$ sera déduite à partir de la table de la loi normale centrée réduite; pour cela on cherche $1,56$ dans la table :
Donc $p(Z \leq 1,56) = 0,9406$.

	...	0.06	...
⋮		⋮	
1,5	...	0,9406	...
⋮			

Pour $94,06\%$ des individus, la variable aléatoire Z est inférieure à $1,56$.

2. Calcul de $p(Z \geq 1,49)$:

$$\begin{aligned} p(Z \geq 1,49) &= 1 - p(Z \leq 1,49) \\ &= 1 - F(1,49). \end{aligned}$$

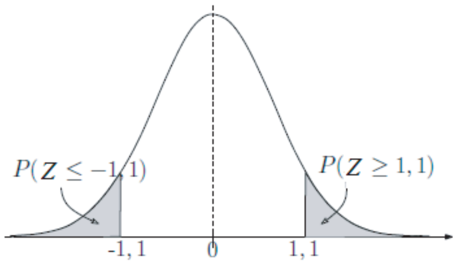
On cherche $1,49$ dans la table :
Donc $p(Z \leq 1,49) = 0,9319$, alors $p(Z \geq 1,49) = 1 - 0,9319 = 0,0681$.

	...	...	0.09
$\vdots$			$\vdots$
$1,4$	...	...	$0,9319$
$\vdots$			

3. Si de plus on veut connaître $p(1,49 \leq Z \leq 1,56$, alors :

$$\begin{aligned} p(1,49 \leq Z \leq 1,56 &= F(1,56) - F(1,49) \\ &= 0,9406 - 0,9319 = 0,0087. \end{aligned}$$

4. On cherche $p(Z \leq -1.1)$, c'est-à-dire $F(-1,1)$.



On sait que $p(Z \leq -1.1) = F(-1,1) = 1 - F(1,1) = 1 - 0,8643 = 0,1357$.
Autrement dit, $p(Z \leq -1,1) = p(Z \geq 1,1) = 1 - 0,8643 = 0,1357$.

Exemple 5.7. Soit X une variable aléatoire continue suivant une loi normale d'une moyenne $m = 2$ et d'un écart type $\sigma = 0,16$.
Quelle est la probabilité d'avoir $1,94 \leq X \leq 2,02$?
Solution

$$X \rightsquigarrow \mathcal{N}(2;0.16) \Rightarrow Z \rightsquigarrow \mathcal{N}(0;1)$$

$$\begin{aligned} p(1,94 \leq X \leq 2,02) &= p(\frac{1,94-2}{0,16} \leq Z \leq \frac{2,02-2}{0,16}) \\ &= p(-0,375 \leq Z \leq 0,125) \\ &= p(Z \leq 0,125) - p(Z \leq -0,375) \\ &= F_Z(0,125) - F_Z(-0,375) = F_Z(0,125) - [1 - F_Z(0,375)] \\ &= 0,5497 - [1 - 0,6462] = 0,1959. \end{aligned}$$

5.2.3 L'approximation de la loi binomiale par une loi de normale

Théorème 5.3. *Soit X une variable aléatoire suivant une loi binomiale de paramètres n et p, i.e. $\mathcal{B}(n;p)$, alors*

$$p(X=k) \simeq \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{k-m}{\sigma})^2} \textit{ quand } n \to +\infty, \textit{ c'est à dire } \mathcal{N}(m;\sigma),$$

avec $m = E(X) = n.p$ *et* $\sigma^2 = V(X) = n.p.q$, *avec* $q = 1 - p$.

Remarque 5.2. On considère que l'approximation est valable si pour $n > 25$, on a à la fois $np > 5$ et $nq > 5$.

En résumé, on a :

$$\begin{array}{lll} \text{Si } n > 25, & & \\ \quad np > 5, & \text{alors la loi binomiale } \mathcal{B}(n;p) \rightsquigarrow & \text{loi Normale } \mathcal{N}(m;\sigma) \\ \quad nq > 5, & & \text{avec } m = np \text{ et } \sigma = \sqrt{n.p.q}. \end{array}$$

Exemple 5.8. Reprenons l'exemple 5.5, en considérant 100 hommes au lieu de 40. rappelons que la probabilité qu'un homme soit atteint d'une maladie M est $p = 0,1$ et la probabilité qu'il ne soit pas atteint de cette maladie est $q = 1 - p = 0,9$. Soit k le nombre d'hommes touchés par la maladie et X la variable aléatoire qui compte le nombre d'hommes malades.

1. Quelle est la loi que suit la v.a. X ? Donner sa loi de distribution $p(X = k)$.
2. Calculer $E(X)$, $V(X)$ et $\sigma(X)$.
3. Par quelle loi peut-on approcher la loi de X ?
4. Calculer $p(X \leq 2)$ et $p(2 \leq X \leq 10)$.

Solution :

1. X suit une loi binomiale $\mathcal{B}(100; 0.1)$:

$$p(X=k) = C_{100}^k.(0,1)^k.(0,9)^{(100-k)}$$

2. Calcul de $E(X)$, $V(X)$ et $\sigma(X)$:

$$\begin{aligned} E(X) &= np = 100 \times 0,1 = 10. \\ V(X) &= npq = 100 \times 0,1 \times 0,9 = 9. \\ \sigma(X) &= \sqrt{V(X)} = 3. \end{aligned}$$

3. Cette loi binomiale $\mathcal{B}(100; 0.1)$ est approchée par une loi de normale $\mathcal{N}(m;\sigma)$:

$$(n = 100 > 25,\ np = 10 > 5 \text{ et } nq = 90 > 5\ ) \Rightarrow \text{binomiale } \mathcal{B}(100;0.1) \rightsquigarrow \text{ normale } \mathcal{N}(m,\sigma),$$

avec $m = np = 10$ et $\sigma = \sqrt{n.p.q} = 3$.

$$p(X = k) \simeq \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{k-m}{\sigma})^2} = \frac{1}{3\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{k-10}{3})^2}.$$

4. Calcul de $p(X \leq 2)$ et $p(2 \leq X \leq 10)$:

$$\begin{aligned}
p(X \leq 2) &= p(\frac{X-m}{\sigma} \leq \frac{2-m}{\sigma}) = p(\frac{X-10}{3} \leq \frac{2-10}{3}) \\
&= p(Z \leq \frac{-8}{3}) = F(\frac{-8}{3}) \text{ où } Z \text{ est la loi normale centrée réduite} \\
&= 1 - F(\frac{8}{3}) = 1 - F(2,66666) \\
&\simeq 1 - 0,980... \simeq 0,02.
\end{aligned}$$

$$\begin{aligned}
p(2 \leq X \leq 10) &= p(\frac{2-m}{\sigma} \leq \frac{X-m}{\sigma} \leq \frac{10-m}{\sigma}) = p(\frac{2-10}{3} \leq Z \leq \frac{10-10}{3}) \\
&= p(Z \leq \frac{10-10}{3}) - p(Z \leq \frac{2-10}{3}) = p(Z \leq \frac{0}{3}) - p(Z \leq \frac{-8}{3}) \\
&= F(0) - F(\frac{-8}{3}) = F(0) - [1 - F(\frac{8}{3})] \\
&= F(0) - [1 - F(2,66666)] \simeq 0,5 - [1 - 0,980...] \\
&\simeq 0,5 - 0,02 = 0,48.
\end{aligned}$$

Remarque. Les valeurs de $F(2,66666)$ et $F(0)$ sont déduites de la table de la loi normale centrée réduite.

5.2.4 L'approximation de la loi de Poisson par une loi normale

Lorsque le paramètre λ d'une loi de Poisson est grand, la loi de Poisson peut être approchée par une loi normale d'espérance λ et de variance λ. Le principe est analogue a celui utilisé pour l'approximation de la loi binomiale par la loi normale.

Théorème 5.4. *Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ, alors*

$$p(X = k) \simeq \frac{1}{\sqrt{\lambda}\sqrt{2\pi}} e^{\frac{-1}{2}(\frac{k-\lambda}{\sqrt{\lambda}})^2} \text{ quand } n \to +\infty,$$

avec $E(X) = \lambda$ et $V(X) = \lambda$.

Remarque 5.3. On considère que l'approximation est valable si $\lambda > 20$.

Si $\lambda > 20$, alors la loi de Poisson $\mathcal{P}(\lambda)$ est approchée par la loi normale $\mathcal{N}(\lambda; \sqrt{\lambda})$.

5.2.5 Table de la loi normale centrée réduite

$$F(t) = P(X \leq t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \text{ et } F(-t) = 1 - F(t).$$

t	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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Utilisation : On lit les décimales dans les lignes et les centièmes en colonnes.
Par exemple, la valeur de $F(1.54)$ se trouve à l'intersection de la ligne 1.5 et de la colonne 0.04 et on trouve $F(1.54) = 0.9382$ à 10^{-4} près.

Bibliographie

[1] A. Ayache and J. Hamonier. *Cours : Statistique Descriptive et Calcul de Probabilités.* Université de Lille. France, 2014.

[2] R. Balan and G. Lamothe. *Une introduction à la biostatistique.* Presses de l'Université du Québec, 2012.

[3] F. Carrat and A. Mallet. *Biostatistique.* Faculté médecine-Université Pierre et Marie Curie, 2013.

[4] M. Colin and G. Payette. *Biostatistiques pour les techniques biologiques.* Montréal, Québec, 3ème edition, 2004.

[5] J.P. Lecoutre. *Statistique et probabilités - Cours et exercices corrigés.* Dunod, 2012.

[6] D. Meghlaoui. *Introduction à la Statistique Descriptive.* Ecole Préparatoire en Sciences Economiques Commerciales et des Sciences de Gestion de Constantine, 2010.

[7] M. MERCIER. *Biostatistique et probabilités.* Ellipses, 2011.

[8] V. Morice and A. Mallet. *QCM corrigées et commentées de Biostatistique.* Ellipses, 2012.

[9] H. Mzali. *Cours : Staitistique et calcul de probabilité.* Ecole Natianale de l'Adminstration. Tunis, 2013.

[10] X. Nogues, A. Garenne, X. Bouteiller, and V. Fiévet. *Le cours de Biostatistique.* Collection tout en fiches : Dunod, 2018.

[11] A. Ruegg. *Probabilités et Statistique.* Presses Polytechniques et universitaires romandes. Suisse, 4ème edition, 1994.

[12] B. Scherrer. *Biostatistique*, volume 1. Gaetan Morin, 2ème edition, Novembre 2008.

[13] B. Scherrer. *Biostatistique*, volume 2. Gaetan Morin, 2ème edition, Mai 2009.

[14] C. Suquet. *Introduction au calcul de probabilités.* Université des Sciences et Technologies de Lille U.F.R. de Mathématiques Pures et Appliquées. France, 2003.