

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique

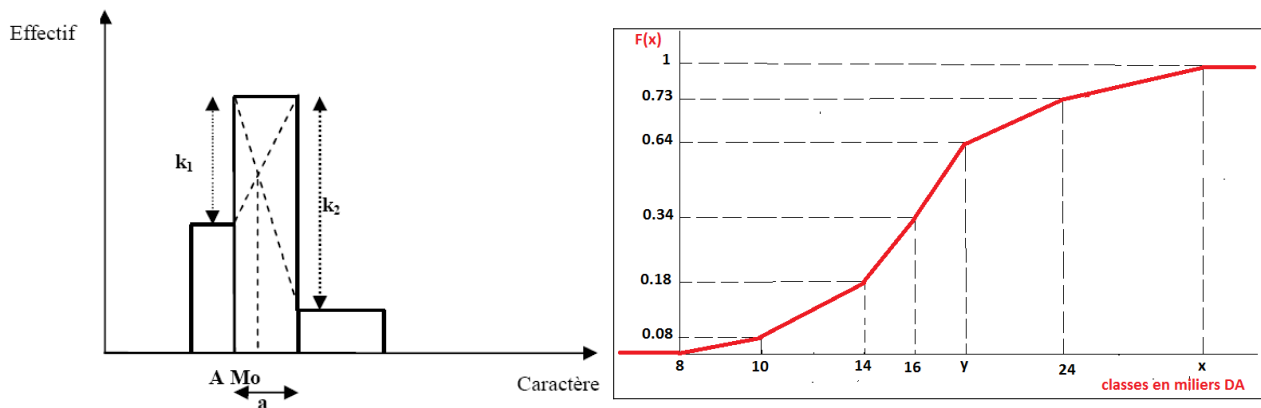
Université 8 Mai 1945 – Guelma  
Faculté des Mathématiques et de l'Informatique  
et des Sciences de la Matière  
Département de Mathématiques



Cours, exercices et examens du module :

# Statistique et Probabilité

(1<sup>ère</sup> LMD MI, 2<sup>ème</sup> LMD ST)



Présenté par: Dr. KERBOUA Mourad

Année Universitaire: 2015/2016

# INTRODUCTION

*La statistique est une discipline qui concerne la quantification de certains phénomènes et l'élaboration de procédures inférentielles. Les méthodes statistiques disponibles actuellement constituent un ensemble de procédures et de règles aidant à l'analyse numérique.*

*On peut faire une distinction entre ces différentes méthodes: Celles relatives à la statistique descriptive (objet de notre module) et celles relatives à la statistique inférentielle.*

*La statistique descriptive est une méthode d'analyse numérique que les scientifiques utilisent pour recueillir, classer, présenter et analyser des informations provenant des ensembles comportant un grand nombre d'éléments.*

*Son utilisation intensive dans de nombreux domaines tels les sciences expérimentales (biologie, physique, agronomie, médecine,...), les sciences humaines, l'économie, etc,...démontre l'intérêt qu'on lui porte.*

*Par contre la statistique inférentielle a pour objectif d'aider à la généralisation de l'information et de faire des inférences: estimation, décision, tests d'hypothèses,..., basées sur des échantillons tirés d'un ou plusieurs univers à étudier.*

# Table des matières

---

<b>1</b>	<b>Concepts de base de la statistique</b>	<b>3</b>
1.1	Introduction . . . . .	3
1.2	Concepts de base de la statistique . . . . .	3
1.2.1	Population et individu . . . . .	3
1.2.2	Variable (ou caractère) . . . . .	4
1.3	Les tableaux statistiques . . . . .	4
1.3.1	Présentation des données sous forme de tableaux . . . . .	4
1.3.2	Règle de STURGES . . . . .	6
1.3.3	Distribution et fonction de répartition . . . . .	6
1.4	Les représentations Graphiques . . . . .	10
1.4.1	Introduction . . . . .	10
1.4.2	Cas de variables qualitatives . . . . .	10
1.4.3	Cas de variables quantitatives . . . . .	11
<b>2</b>	<b>Représentation numérique des données</b>	<b>15</b>
2.1	Les Caractéristiques de tendance centrale ou de position . . . . .	15
2.1.1	<b>La Médiane</b> . . . . .	15
2.1.2	<b>Les quartiles</b> . . . . .	18
2.1.3	<b>Intervalle interquartile</b> . . . . .	20
2.1.4	<b>Le mode</b> . . . . .	20
2.1.5	<b>La moyenne arithmétique</b> . . . . .	22
2.1.6	<b>La moyenne arithmétique pondérée</b> . . . . .	22
2.1.7	<b>La moyenne géométrique</b> . . . . .	23
2.1.8	<b>La moyenne harmonique</b> . . . . .	24

2.1.9	<b>La moyenne quadratique</b>	24
2.2	Les caractéristiques de dispersion	25
2.2.1	<b>L'étendu</b>	25
2.2.2	<b>La variance</b>	25
2.2.3	<b>L'écart type</b>	26
2.2.4	<b>L'écart absolue moyen</b>	26
2.2.5	<b>Le coefficient de variation</b>	26
<b>3</b>	<b>Calculs des probabilités</b>	<b>28</b>
3.1	Analyse combinatoire	28
3.1.1	Principe fondamental de l'analyse combinatoire	28
3.1.2	Arrangements	28
3.1.3	Permutations	30
3.1.4	Combinaisons	31
3.2	Calculs des probabilités	32
3.2.1	Espace probabilisable	32
3.2.2	Lien entre la théorie des probabilités et des ensembles	35
3.2.3	Construction d'une probabilité	36
3.2.4	Probabilités conditionnelles, indépendance et probabilités composées	37
<b>4</b>	<b>TD et EXAMENS</b>	<b>41</b>
4.1	Série de TD 1	41
4.2	Série de TD 3 (Analyse Combinatoire)	43
4.3	Série de TD 4 (Calcul des probabilités)	45
4.4	Examen 2012	49
4.5	Examen de Rattrapage 2012	54
4.6	Examen 2011	59
4.7	Examen de Rattrapage 2011	61
4.8	Examen 2010	65
4.9	Examen de Rattrapage 2010	67
4.10	correction examen de Rattrapage 2010	69
<b>5</b>	<b>Exercices Divers</b>	<b>73</b>

# Concepts de base de la statistique

---

## 1.1 Introduction

*La statistique est une discipline qui concerne la quantification de certains phénomènes et l'élaboration de procédures inférentielles. Les méthodes statistiques disponibles actuellement constituent un ensemble de procédures et de règles aidant à l'analyse numérique.*

*On peut faire une distinction entre ces différentes méthodes : Celles relatives à la statistique descriptive (objet de notre module) et celles relatives à la statistique inférentielle.*

*La statistique descriptive est une méthode d'analyse numérique que les scientifiques utilisent pour recueillir, classer, présenter et analyser des informations provenant des ensembles comportant un grand nombre d'éléments.*

*Son utilisation intensive dans de nombreux domaines tels les sciences expérimentales (biologie, physique, agronomie, médecine,...), les sciences humaines, l'économie, etc,...démontre l'intérêt qu'on lui porte.*

*Par contre la statistique inférentielle a pour objectif d'aider à la généralisation de l'information et de faire des inférences : estimation, décision, tests d'hypothèses,..., basées sur des échantillons tirés d'un ou plusieurs univers à étudier.*

## 1.2 Concepts de base de la statistique

Dans une manière générale, la méthode statistique est basée sur quatre concepts : la population, les variables, les observations et les données.

### 1.2.1 Population et individu

La notion fondamentale en statistique est celle d'ensemble d'objets équivalents que l'on appelle *population*, ses objets sont appelés *individus* ou *unités statistiques*.

Lorsque l'on observe qu'une partie de la population, on parle de *sondage*.

La partie de la population étudiée est appelée *échantillon* et on cherche toujours à généraliser les résultats obtenus sur l'échantillon à toute la population.

Exemple 1

- ▶ Etudier les notes au module de statistique sur un échantillon de 1000 étudiants est une expérience statistique.  
Les unités statistique ou bien les individus correspondent aux étudiants.  
La population est l'ensemble des étudiants.
- ▶ Lors de l'étude sur l'emploi, la population pourrait être l'ensemble des personnes en âge de travailler.

### 1.2.2 Variable (ou caractère)

Chaque individu d'une population est décrit par un ensemble de caractéristiques appelés *variables* ou *caractères*. On les représente souvent par des lettres majuscules :  $X, Y, \dots$

Les valeurs que peut prendre une variable statistique sont appelées *modalités*.

Une variable doit donc présenter au minimum deux modalités.

Exemple 2

- ▶ La variable «sexe» a deux modalités : masculin ou féminin.
- ▶ Les modalités de la variable «âge des ouvriers d'une entreprise» peuvent être :  $[25, 30[$ ;  $[30, 40[$ ;  $[40, 45[$ ;  $[45, 60[$ .

Les variables statistiques peuvent être classées selon leur nature en deux catégories :

#### a) Variable quantitative

Une variable est dite «quantitative» si ces différentes valeurs sont mesurables.

Il y a deux types de variables quantitatives :

##### ◆ Variable discrète

Une variable quantitative est dite *discrète* si elle prend un nombre fini ou dénombrable de valeurs.

Exemple 3

La variable nombre d'enfants par famille, dont les valeurs sont : 0, 1, 2, 3, etc,... est une variable quantitative discrète.

##### ◆ Variable continue

Une variable quantitative est dite *continue* si elle prend toutes ses valeurs dans un intervalle de l'ensemble des nombre réels  $\mathbb{R}$ .

Exemple 4

La variable «taille en cm» observée sur un échantillon de 500 individus est une variable quantitative continue, dont les résultats sont les suivants :

100 individus leurs tailles sont comprise dans l'intervalle  $[150, 160[$ , 250 individus dans l'intervalle  $[160, 170[$  et 150 dans l'intervalle  $[170, 185[$ .

#### b) Variable qualitative

Une variable est dite *qualitative* si ses différents modalités ne sont pas mesurables.

Exemple 5

Le sexe, la nationalité, la couleur des yeux, la catégorie socioprofessionnelle.

## 1.3 Les tableaux statistiques

### 1.3.1 Présentation des données sous forme de tableaux

Au cours d'une étude statistique les données sont recueillies de façon désordonnée. Cette masse d'informations que si l'information en question est classées et mises en ordre, ce classement ne peut se faire que dans des tableaux statistiques qui servent de documentation statistique.

Leur présentation diffère selon la nature des variables.

**a) Cas d'une variable discrète :**

Lorsque la variable statistique est discrète, il arrive souvent qu'une même valeur  $x_i$  ( $i = 1, \dots, k$ ;  $k =$  nombre de modalités) soit observée sur plusieurs individus, on note alors  $n_i$  son effectif (fréquence absolue), qui est le nombre d'observations égales à  $x_i$  et  $f_i$  la fréquence relative correspondantes :  $f_i = \frac{n_i}{n}$ ,  $i = 1, 2, \dots, k$ .

Le tableau statistique se présente alors sous la forme :

$x_i$	$n_i$	$f_i$

**Remarque 1**

$$\sum_{i=1}^k n_i = n, \quad \sum_{i=1}^k f_i = 1$$

*Exemple 1*

Une étude sur la répartition des ménages selon le nombre de personnes qui les composent nous donne le tableau suivant :

Nombre de personne dans un ménage $x_i$	Nombre de ménages $n_i$
1	50
2	80
3	90
4	120
5	350
6	160
7 et plus	150
Total	1000

**b) Cas d'une variable continue**

Lorsque la variable est continue, on regroupe les observations en  $k$ -classes d'extrémités  $c_0, c_1, \dots, c_k$ .

On note par chaque classe  $[c_i, c_{i+1}[$  l'effectif  $n_i$  et la fréquence relative  $f_i$  le tableau statistique se présente alors sous la forme suivante :

$[c_i, c_{i+1}[$	$n_i$	$f_i$

Par convention la borne supérieure d'une classe est exclue de cette classe, sauf pour (la dernière classe peut être fermée).

On appelle «centre» d'une classe  $[c_i, c_{i+1}[$  la valeur  $y_i = \frac{c_i + c_{i+1}}{2}$  et «l'amplitude» de cette classe est notée  $a_i$ , telle que :  $a_i = c_{i+1} - c_i$ .

*Exemple 2*

On a observés les tailles (en cm) de 24 personnes, les résultats sont les suivant : **170, 165, 170, 165, 175, 165, 178, 178, 162, 159, 158, 176, 180, 180, 182, 179, 183, 165, 176, 177, 180, 173, 173, 162.**

Ces observations peuvent être représenté par des classes d'amplitude 4 cm par exemple :

Tailles	Effectifs partiels $n_i$	Fréquences $f_i$	Fréquences en pourcentage
[158 – 162[	2	0.083	8.3%
[162 – 166[	6	0.25	25%
[166 – 170[	0	0	0%
[170 – 174[	4	0.167	16.7%
[174 – 178[	4	0.167	16.7%
[178 – 182[	6	0.25	25%
[182 – 186[	2	0.083	8.3%
<b>Total</b>	24	1	100%

### 1.3.2 Règle de STURGES

Cette règle est utilisée pour déterminer le nombre de classes à utiliser pour représenter une variable statistique continue.

Le nombre de classes est égal à l'entier le plus proche de la quantité :  $1 + 3.3 \log_{10} n$ .

Par exemple, pour un échantillon de taille  $n = 200$ , on doit utiliser

$1 + 3.3 \log_{10} 200 = 1 + 3.3 \times 2.3 = 8.59 \approx 9$  classes

L'amplitude constante de ces classes sera égale à :  $a = \frac{X_{\max} - X_{\min}}{\text{Nombre de classes}}$

### 1.3.3 Distribution et fonction de répartition

#### Distribution

Les distribution d'une variable statistique  $X$  est l'ensemble des couples  $(x_i, f_i)$ ,  $i = 1, 2, \dots, k$  ( $k =$  nombre de modalités de la variable étudiée)

*Exemple 3*

$(0; 0.2), (1; 0.3), (2; 0.4), (3 \text{ et plus}; 0.1)$  est une distribution statistique représentant la variable «nombre d'enfants par famille».

#### Fonction de répartition

La fonction de répartition  $F(x)$ , appelée aussi fonction des fréquences cumulées, associé à tout nombre réel  $x$  la proportion des individus pour lesquels la variable statistique  $X$  prend des valeurs inférieure à  $x$ , donc :

$$F(x) = \sum_{x_i < x} f_i$$

*Exemple 4*

Les taux de glécimie, observés sur un échantillon de 100 patients sont donnés dans le tableau suivant :



Taux de glécimie	$[0.8 - 0.85[$	$[0.85 - 0.9[$	$[0.9 - 0.95[$	$[0.95 - 1[$	$[1 - 1.20[$	Total
Effectifs $n_i$	10	20	30	30	10	100
Fréquences $f_i$	0.1	0.2	0.3	0.3	0.1	1
$F(x)$	0	□ 0.1	□ 0.3	□ 0.6	□ 0.9	1 □

Pour :

$$x = 0.8; \quad F(0.8) = \sum_{x_i < 0.8} f_i = 0$$

$$x = 0.85; \quad F(0.85) = \sum_{x_i < 0.85} f_i = f_1 = 0.1$$

$$x = 0.9; \quad F(0.9) = \sum_{x_i < 0.9} f_i = f_1 + f_2 = 0.3$$

$$x = 0.95; \quad F(0.95) = \sum_{x_i < 0.95} f_i = f_1 + f_2 + f_3 = 0.6$$

$$x = 1; \quad F(1) = \sum_{x_i < 1} f_i = f_1 + f_2 + f_3 + f_4 = 0.9$$

$$x = 1.20; \quad F(1.20) = \sum_{x_i < 1.20} f_i = f_1 + f_2 + f_3 + f_4 + f_5 = 1$$

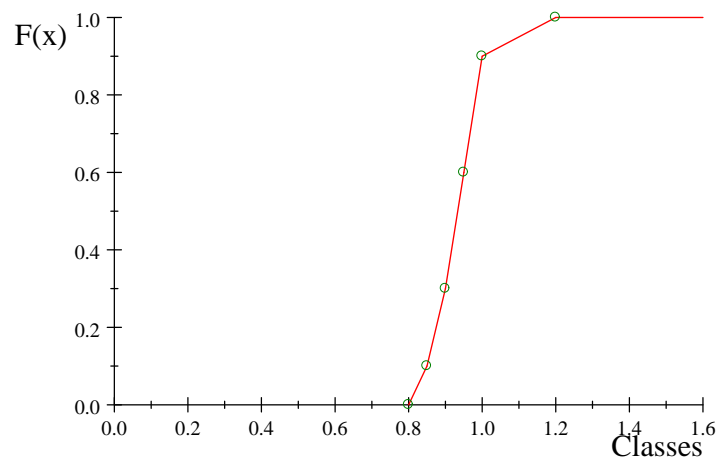
$F(0.95) = 0.6$  par exemple, veut dire que 60% des patients ont un taux de glécimie inférieur à 0.95.

### Remarque 2

► On a  $F(x) = 0$ , si  $x \leq \min x_i$  et  $F(x) = 1$ , si  $x > \max x_i$ .

►  $F$  est une fonction monotone croissante, elle prend ses valeurs dans l'intervalle  $[0, 1]$  et présente lorsque  $X$  est discrète une discontinuité en chacun des points  $x_i$  qui se manifeste par un saut dont la valeur est  $f_i$ .

Ainsi la fonction de répartition précédente peut être représenté graphiquement par une courbe continue (courbe cumulative) :

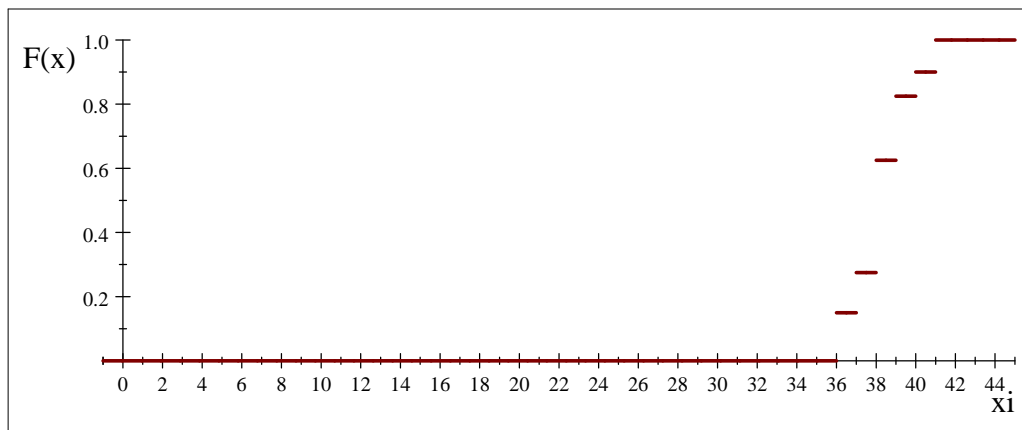


### Exemple 5

Une enquête faite sur les pointures de 200 personnes a fourni les résultats suivants : Comme la variable étudiée est discrète, alors la fonction de répartition va être représenté par une courbe en escaliers :

$x_i$	36	37	38	39	40	41	Total
$n_i$	30	25	70	40	15	20	200
Fréquences $f_i$	0.15	0.125	0.35	0.2	0.075	0.1	1
$F(x)$	0	0.15	0.275	0.625	0.825	0.9	1

$$F(x) = \begin{cases} 0 & \text{si } x < 36 \\ 0.15 & \text{si } 36 \leq x < 37 \\ 0.275 & \text{si } 37 \leq x < 38 \\ 0.625 & \text{si } 38 \leq x < 39 \\ 0.825 & \text{si } 39 \leq x < 40 \\ 0.9 & \text{si } 40 \leq x < 41 \\ 1 & \text{si } x \geq 41 \end{cases}$$



### Courbes cumulatives croissante et décroissante

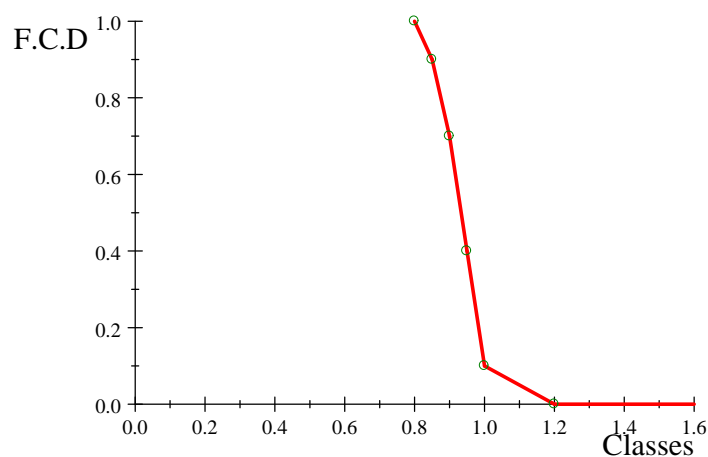
Pour définir la fonction de répartition, on calcule les fréquences cumulées croissantes, qu'on représente graphiquement par une courbe cumulative croissante.

De même façon on peut calculer aussi les fréquences cumulées décroissantes, à qui correspond une courbe cumulative décroissante, notée :

$$F(x) \downarrow = \sum_{x_i \geq x} f_i$$

Calculons les fréquences cumulées décroissantes de la variable étudiée dans l'exemple 4 :

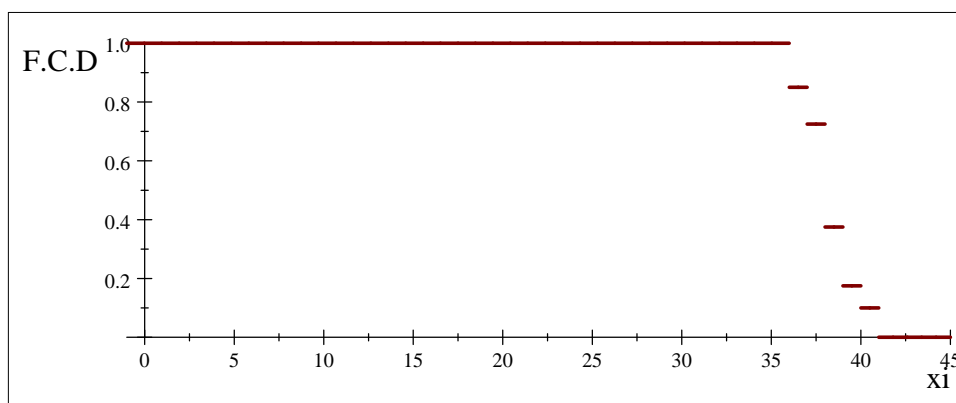
Taux de glécimie	[0.8 – 0.85[	[0.85 – 0.9[	[0.9 – 0.95[	[0.95 – 1[	[1 – 1.20[	Total
Fréquences $f_i$	0.1	0.2	0.3	0.3	0.1	1
$F(x) \downarrow$	1	0.9	0.7	0.4	0.1	0 /
$F(x)$	0	0.1	0.3	0.6	0.9	1 /



Courbe cumulative décroissante

$F(0.9) \downarrow = 0.7$ , c'est-à-dire 70% des patients ont un taux de glécimie  $\geq 0.9$ .  
 $F(0.9) \uparrow = 0.3$ , c'est-à-dire 30% des patients ont un taux de glécimie  $< 0.9$ .  
 Revenons à l'exemple 5.

$x_i$	36	37	38	39	40	41	Total
Fréquences $f_i$	0.15	0.125	0.35	0.2	0.075	0.1	1
$F(x) \downarrow$	1	0.85	0.725	0.375	0.175	0.1	0 /



**Remarque 3**

Les courbes cumulative croissante et décroissante peuvent être obtenues aussi à l'aides des effectifs cumulés croissant (noté  $N_i \uparrow$ ) et les effectifs cumulés décroissant (noté  $N_i \downarrow$ ).

## 1.4 Les représentations Graphiques

### 1.4.1 Introduction

*Bien qu'un tableau statistique contient toute l'information prélevé sur l'échantillon, mais il est très utile de traduire ces informations par un graphique, a fin d'avoir un synthèse visuelle. Suivant la nature de la variable étudiée, on utilise différents types de représentations.*

### 1.4.2 Cas de variables qualitatives

#### Représentation circulaire par des secteurs

Elle consiste à représenter sur un cercle chaque modalité par un secteur dont le degré est déterminé à l'aide de la règle de trois de la manière suivante :

Supposons qu'on veut calculer le degré du secteur représentant la modalité  $i$  d'effectif partiel  $n_i$  alors on peut écrire :

$$\left\{ \begin{array}{l} n \text{ (l'effectif total)} \longrightarrow 360^\circ \\ \implies d_i = \frac{n_i \times 360^\circ}{n} = f_i \times 360^\circ, \quad i = 1, \dots, k \\ \qquad \qquad \qquad n_i \qquad \qquad \qquad \longrightarrow d_i \end{array} \right.$$

#### *Exemple 1*

On a étudié le régime des études sur un échantillon de 200 élèves d'un lycée donné, les résultats obtenus sont les suivants :

Régime des études	Externe	Interne	Demi-pensionnaire
Nombre d'élèves $n_i$	70	50	80
Fréquences $f_i$	0.35	0.25	0.4

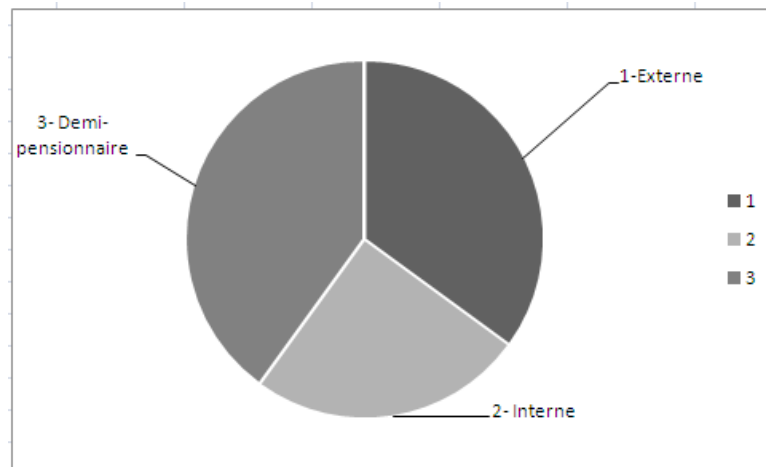
calculons le degré correspondant aux différents secteurs :

$$d_1 = f_1 \times 360^\circ = 0.35 \times 360^\circ = 126^\circ$$

$$d_2 = f_2 \times 360^\circ = 0.25 \times 360^\circ = 90^\circ$$

$$d_3 = f_3 \times 360^\circ = 0.4 \times 360^\circ = 144^\circ$$

D'où on obtient la représentation suivante :



### 1.4.3 Cas de variables quantitatives

On va distinguer les diagrammes correspondant à une variable discrète et ceux correspondant à une variable continue.

#### Le diagramme en bâtons

Il consiste à représenter les divers valeurs  $x_i$  prises par une variable discrète, en fonction des fréquences  $f_i$  ou des effectifs  $n_i$  de la manière suivante :

On représente en abscisse les observations  $x_i$  et au dessus de chaque valeur on trace un bâton dont la hauteur est proportionnelle à  $f_i$  ou  $n_i$ .

#### Exemple 2

Les poids (en kg) de 200 individus sont représentés dans le tableau suivant :

$x_i$	55	60	65	70	75	80	<b>Total</b>
$n_i$	20	60	40	20	50	10	200
$f_i$	0.1	0.3	0.2	0.1	0.25	0.05	1

Cette variable est discrète, donc on peut la représenter graphiquement par le diagramme en bâtons en utilisant les fréquences  $f_i$  ou les effectifs  $n_i$ .

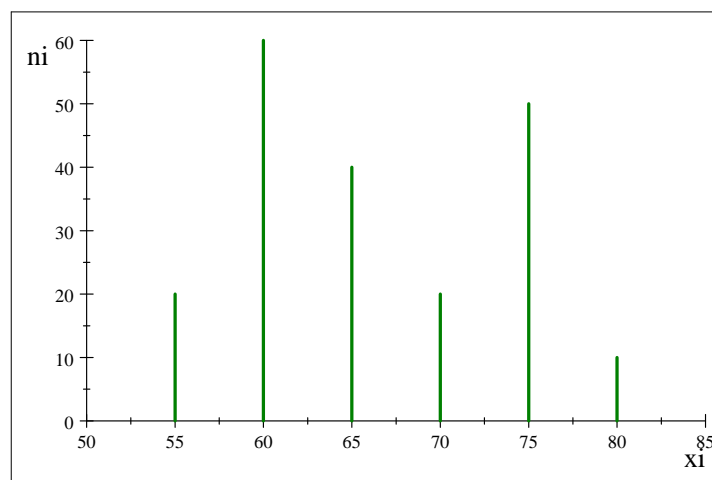


Diagramme en bâtons

### Histogramme

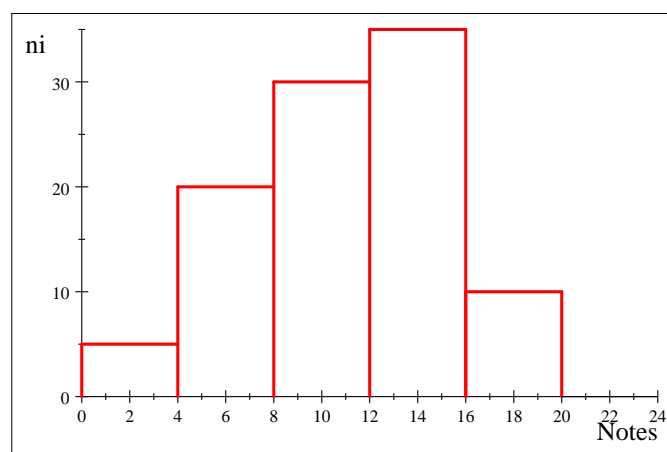
L'histogramme est utilisé pour représenter une variable continue. On représente sur l'axe des abscisses les différentes classes (supposée de même amplitude) et au dessus de chacune d'elle on trace des rectangles, dont les hauteurs sont directement proportionnelles aux fréquences ou aux effectifs.

#### Exemple 3

Les notes au module de statistique, observées sur un échantillon de 100 étudiants sont données dans le tableau suivant :

Notes	[0 – 4[	[4 – 8[	[8 – 12[	[12 – 16[	[16 – 20]	<b>Total</b>
Effectifs $n_i$	5	20	30	35	10	100

La variable étudiée est continue, donc elle peut être représenté graphiquement au moyen de l'histogramme suivant :



Histogramme

**Remarque 1**

Ce mode de représentation est valable seulement dans le cas où on a des classes d'intervalles d'amplitudes constantes.

Dans le cas où les amplitudes sont différentes, pour tracer l'histogramme on utilise la méthode des effectifs rectifiés, qui consiste à multiplier l'effectif de chaque classe par le rapport entre l'amplitude de la plus petite classe sur l'amplitude de la classe considérée, on obtient ainsi les effectifs rectifiés noté  $n'_i$  :

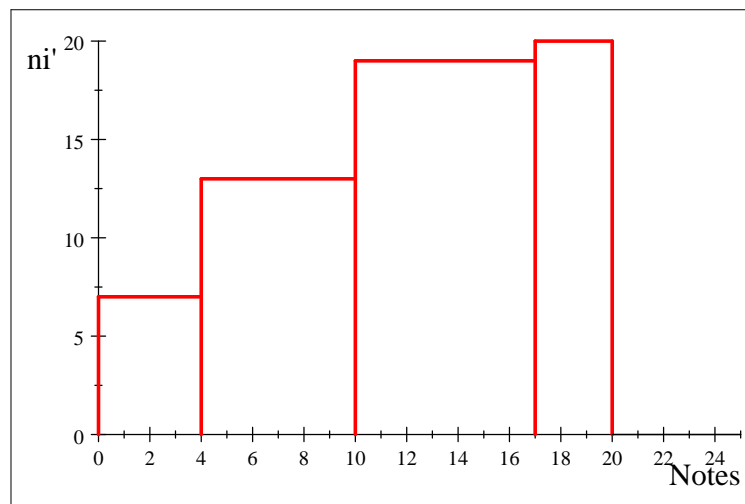
$$n'_i = n_i \times \frac{\text{la plus petite amplitude}}{\text{amplitude de la classe } i}$$

Exemple 4

On reprend l'exemple précédent avec une nouvelle distribution.

Notes	$[0 - 4[$	$[4 - 10[$	$[10 - 17[$	$[17 - 20]$	<b>Total</b>
Effectifs $n_i$	10	25	45	20	100
Amplitudes $a_i$	4	6	7	3	/
Effectifs rectifiés $n'_i$	$10 \times \frac{3}{4} \approx 7$	$25 \times \frac{3}{6} \approx 13$	$45 \times \frac{3}{7} \approx 19$	$20 \times \frac{3}{3} = 20$	/

D'où l'histogramme correspondant :

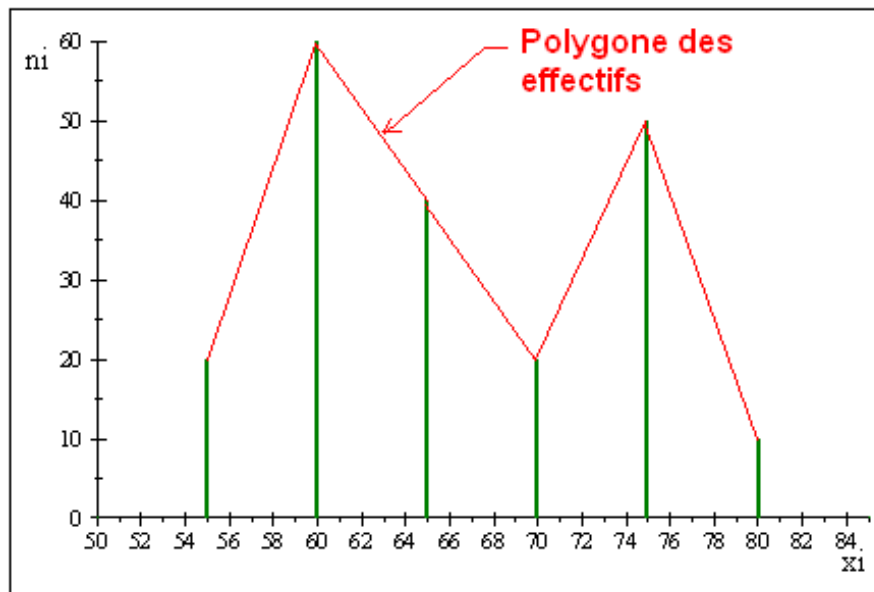
**Remarque 2**

La surface de chaque rectangle est égale à la fréquence de la classe correspondante ( $s_i = \frac{f_i}{a_i} \times a_i = f_i$ ), donc la surface totale de l'histogramme est égale à 1.

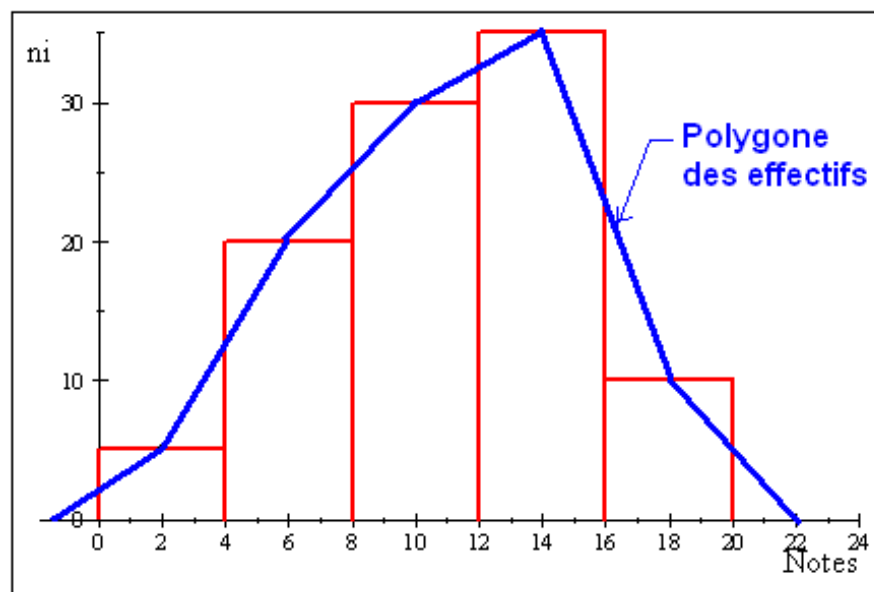
**Polygone**

Le polygone est construit à partir du diagramme en batons ou de l'histogramme. Il s'agit d'une ligne brisée qui joint les points supérieurs des batons ou les points milieux des côtés supérieurs des rectangles successifs de l'histogramme.

Le polygone de effectifs correspondants à la variable étudiée dans l'exemple 2 :



Le polygone des effectifs correspondant à la variable étudiée dans l'exemple 3 :





# Représentation numérique des données

---

Une série de données peut être résumée par quelques valeurs numériques appelées *caractéristiques des séries statistiques*, classées en trois grandes catégories :

- les caractéristiques de tendance centrale,
- les caractéristiques de dispersion,
- les caractéristiques de forme,

## 2.1 Les Caractéristiques de tendance centrale ou de position

Elles donnent une idée de l'ordre de grandeur des valeurs constituant la série ainsi que la position où semblent se concentrer les valeurs de cette série. Les principales caractéristiques de tendance centrale sont la *médiane*, Les *quartiles*, Le *mode*, La *moyenne* (arithmétique, géométrique, harmonique, quadratique)

### 2.1.1 La Médiane

On appelle "**médiane**" de la variable statistique  $X$ , tout nombre réel  $Me$  tel que  $F(Me) = 0.5$ .

C'est à dire, la proportion des individus pour les quels  $X$  prend des valeurs plus petite (respectivement plus grande) que  $Me$ , ne dépasse pas 50 %.

#### ► Détermination de $Me$ dans le cas d'une variable discrète :

On doit d'abord ordonner la série statistique dans l'ordre croissant :

$x_1 \leq x_2 \leq \dots \leq x_{n-1} \leq x_n$ . Par la suite on peut avoir deux cas :

1<sup>er</sup> cas : Si  $n$  est un nombre impair alors on peut écrire  $n = 2p + 1$ , avec  $p$  un entier.

Dans ce cas  $Me = x_{p+1}$  où  $x_{p+1}$  est l'observation numéro  $(p + 1)$  dans l'ordre croissant.

– **Exemples** :

1) Les notes obtenues par 15 étudiants au module de mathématiques sont :  
14, 16, 8, 10, 17, 11, 10, 8, 13, 14, 9, 7, 15, 10, 12.

Ordonnons les observations dans l'ordre croissant

7 , 8, 8, 9, 10, 10, 10, 11 , 12, 13, 14, 14, 15, 16, 17  
 $\downarrow$   $\downarrow$   $\downarrow$   
 $x_1$   $x_8$   $x_{15}$

$n = 15$  est un nombre impair alors  $n = 2 \times 7 + 1 \Rightarrow Me = x_8 = 11$ .

2) On considère la même variable, mais observé sur un échantillon de 121 étudiants

$x_i$	6	8	9	10	12	13	15	Total
$n_i$	10	20	15	30	25	11	10	121

Ordonnons la série statistique dans l'ordre croissant :

$\underbrace{6 \dots 6}_{10}$   $\underbrace{8 \dots 8}_{20}$   $\underbrace{9 \dots 9}_{15}$   $\underbrace{10 \dots 10}_{30}$   $\underbrace{12 \dots 12}_{25}$   $\underbrace{13 \dots 13}_{11}$   $\underbrace{15 \dots 15}_{10}$

$n = 121 = 2 \times 60 + 1$ , donc  $p = 60 \Rightarrow Me = x_{61} = 10$ .

La valeur qui partage les observations en deux parties égales est la médiane  $Me = 11$ .

2<sup>ème</sup> cas : Si  $n$  est un nombre pair alors on peut écrire  $n = 2p$ , avec  $p$  un entier. Dans ce cas  $Me = \frac{x_p + x_{p+1}}{2}$ .

– **Exemples** :

1) Les notes obtenues par 10 étudiants au module de physique sont :

12, 10, 8, 9, 15, 14, 7, 11, 15, 8.

Ordonnons la série statistique dans l'ordre croissant :

7 , 8, 8, 9, 10 , 11 , 12, 14, 15, 15  
 $\downarrow$   $\downarrow$   $\downarrow$   $\downarrow$   
 $x_1$   $x_5$   $x_6$   $x_{10}$

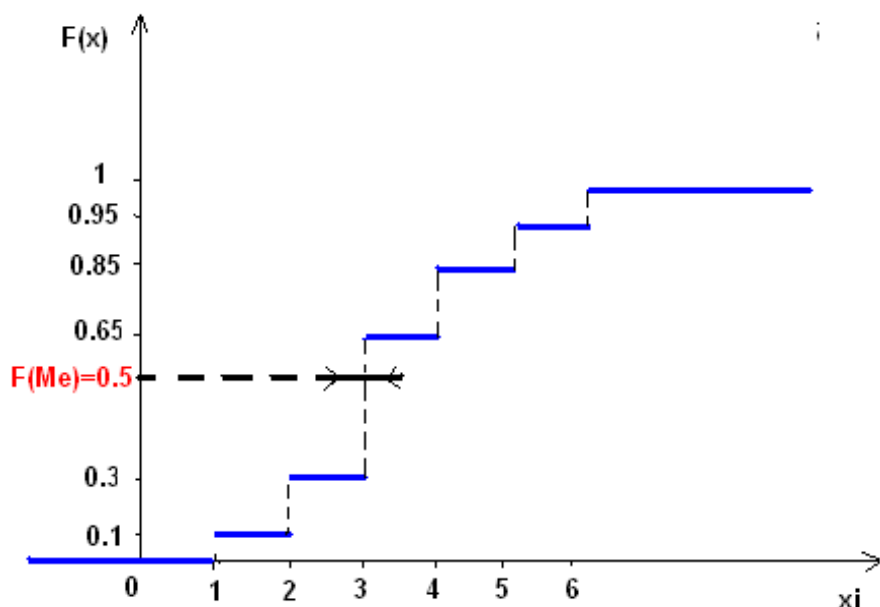
$n = 10$  est nombre pair alors  $n = 2 \times 5$ , avec  $p = 5 \Rightarrow Me = \frac{x_5 + x_6}{2} = \frac{10 + 11}{2} = 10.5$ .

Donc 50 % des étudiants observés ont eu des notes inférieure à 10.5.

On peut aussi déterminer la médiane graphiquement à partir de la courbe cumulative croissante ou décroissante.

2) Une enquête effectuée sur la répartition de 200 foyers selon le nombre de chambre a donné les résultats suivant :

Nbre de chambre $x_i$	1	2	3	4	5	6	Total
$n_i$	20	40	70	40	20	10	200
$f_i$	0.1	0.2	0.35	0.2	0.1	0.05	1
$F(x)$	0	0.1	0.3	0.65	0.85	0.95	1



D'après la courbe cumulative la valeur  $F(Me) = 0.5$  correspond à l'observation 3, d'où la médiane  $Me = 3$ .

► **Détermination de  $Me$  dans le cas d'une variable continue :**

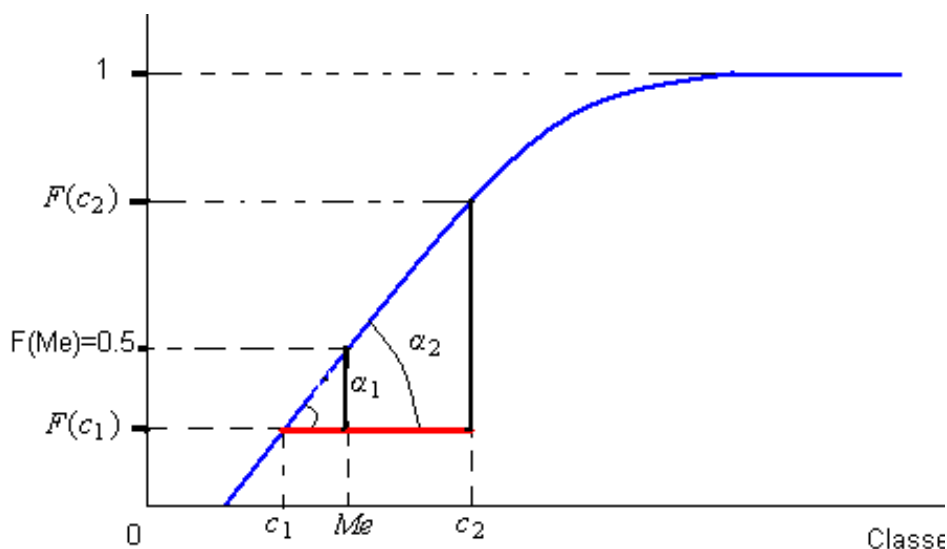
Pour déterminer la médiane d'une variable continue, on doit d'abord la localiser, c'est-à-dire déterminer la classe qui contienne et ceci à partir des fréquences cumulées de la manière suivante :

Si la valeur 0.5 est comprise entre  $F(c_1)$  et  $F(c_2)$  alors la médiane  $Me$  doit se trouver forcément dans la classe  $[c_1, c_2[$ , c'est-à-dire.

$$F(c_1) \leq 0.5 < F(c_2) \implies Me \in [c_1, c_2[.$$

Cette classe est appelée classe médiane.

Une fois que la médiane est localisée, on utilise la méthode d'Interpolation linéaire pour trouver une approchée :



$$\text{On a } \alpha_1 \approx \alpha_2 \implies \tan \alpha_1 \approx \tan \alpha_2 \implies \frac{0.5 - F(c_1)}{Me - c_1} \approx \frac{F(c_2) - F(c_1)}{c_2 - c_1}$$

$$\implies \boxed{Me \approx c_1 + (c_2 - c_1) \frac{0.5 - F(c_1)}{F(c_2) - F(c_1)}}$$

– **Exemple**

Un échantillon de 100 mesures (en mm) de pièces de rechange, fabriquées dans une entreprise a donné les résultats suivants :

Mesures	[40, 45[	[45, 50[	[50, 55[	[55, 60[	[60, 65[	Total
Effectifs $n_i$	30	10	20	25	15	100
Fréquences $f_i$	0.3	0.1	0.2	0.25	0.15	1
Fréquences cumulées $F(x)$	0	0.3	0.4	0.6	0.85	1

↔

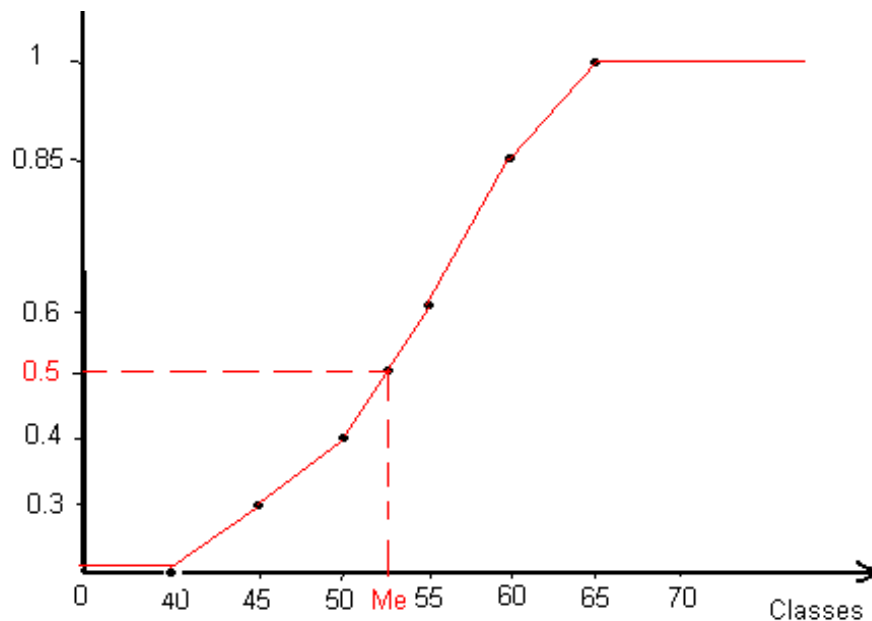
0.5

On a  $F(50) \leq 0.5 < F(55) \implies Me \in [50, 55[$ , donc la classe médiane est la classe  $[50, 55[$ .

De là

$$Me \approx c_1 + (c_2 - c_1) \frac{0.5 - F(c_1)}{F(c_2) - F(c_1)} \approx 50 + (55 - 50) \frac{0.5 - 0.4}{0.6 - 0.4} \approx 52.5$$

On peut aussi déterminer la médiane graphiquement.

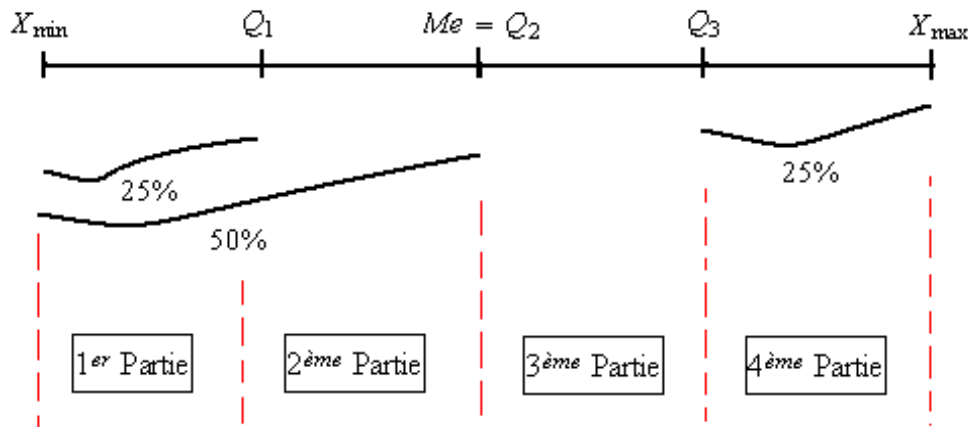


D'après la courbe, la valeur  $F(Me) = 0.5$  correspond à l'observation 52.5, alors la médiane  $Me \approx 52.5$ .

### 2.1.2 Les quartiles

C'est la généralisation de la notion de la médiane, ils s'agissent des valeurs qui partagent les observations en 4 parties égales :

Le premier quartile noté  $Q_1$ , tel que  $F(Q_1) = 0.25$ , le deuxième quartile noté  $Q_2$  est en même temps la médiane, tel que  $F(Q_2) = 0.5$  et le troisième quartile noté  $Q_3$ , tel que  $F(Q_3) = 0.75$ .



► Calcul des quartiles dans le cas d'une variable discrète :

Par analogie au calcul de la médiane vu précédemment, on ordonne d'abord les observations dans l'ordre croissant, deux cas peuvent se présenter :

**1<sup>er</sup> Cas :** Si  $n$  est un nombre pair alors on peut écrire  $n = 2 \times p$  et

$$Q_1 = \frac{x_{\frac{p}{2}} + x_{\frac{p}{2}+1}}{2}, \quad Q_2 = \frac{x_p + x_{p+1}}{2}, \quad Q_3 = \frac{x_{\frac{3p}{2}} + x_{\frac{3p}{2}+1}}{2}$$

**2<sup>ème</sup> Cas :** Si  $n$  est un nombre impair alors on peut écrire  $n = 2 \times p + 1$  et

$$Q_1 = x_{\frac{p}{2}+1}, \quad Q_2 = x_{p+1}, \quad Q_3 = x_{\frac{3p}{2}+1}$$

► Calcul les quartiles dans le cas d'une variable continue

La méthode d'interpolation utilisée pour le calcul de la médiane est appliquée dans ce cas pour déterminer les différents quartiles.

Supposons que les différents quartiles sont localisés respectivement dans les classes  $[c_1, c_2[$ ,  $[c_2, c_3[$  et  $[c_3, c_4[$  alors

$$Q_1 = c_1 + (c_2 - c_1) \frac{0.25 - F(c_1)}{F(c_2) - F(c_1)}$$

$$Q_2 = c_2 + (c_3 - c_2) \frac{0.5 - F(c_2)}{F(c_3) - F(c_2)}$$

$$Q_3 = c_3 + (c_4 - c_3) \frac{0.75 - F(c_3)}{F(c_4) - F(c_3)}$$

◇ **Remarque 1**

*Les quartiles peuvent être déterminés aussi graphiquement à l'aide de la courbe cumulative.*

Notes	[0 – 5[	[5 – 10[	[10 – 15[	[15 – 20[	Total
Effectifs $n_i$	15	20	45	20	100
Fréquences $f_i$	0.15	0.2	0.45	0.2	1
Fréquences cumulées $F(x)$	0	0.15	0.35	0.8	1

### 2.1.3 Intervalle interquartile

L'intervalle interquartile est la différence entre les valeurs du 3<sup>ème</sup> et du 1<sup>er</sup> quartile :  $Q_3 - Q_1$ .

Autrement dit, c'est l'intervalle qui regroupe 50 % des observations, en laissant 25 % à droite et 25 % à gauche.

#### – Exemple

Les notes au module de statistique descriptive obtenues par 100 étudiants sont données dans le tableau suivant :

Le premier quartile  $Q_1$  : (localisation d'abord)

$$\underbrace{F(5)}_{=0.15} \leq 0.25 \leq \underbrace{F(10)}_{=0.35} \implies Q_1 \in [5, 10[, \text{ alors}$$

$$Q_1 = 5 + (10 - 5) \frac{0.25 - 0.15}{0.35 - 0.15} = 7.5$$

C'est-à-dire, 25 % des étudiants ont obtenus des notes inférieure à 7.5.

Le deuxième quartile

$$\underbrace{F(10)}_{=0.35} \leq 0.5 \leq \underbrace{F(15)}_{=0.8} \implies Q_2 \in [10, 15[, \text{ alors}$$

$$Q_2 = 10 + (15 - 10) \frac{0.5 - 0.35}{0.8 - 0.35} = 11.67$$

C'est-à-dire, 50 % des étudiants ont obtenus des notes inférieure à 11.67.

Le troisième quartile

$$\underbrace{F(10)}_{=0.35} \leq 0.75 \leq \underbrace{F(15)}_{=0.8} \implies Q_3 \in [10, 15[, \text{ alors}$$

$$Q_3 = 10 + (15 - 10) \frac{0.75 - 0.35}{0.8 - 0.35} = 14.44$$

C'est-à-dire, 75 % des étudiants ont obtenus des notes inférieure à 14.44.

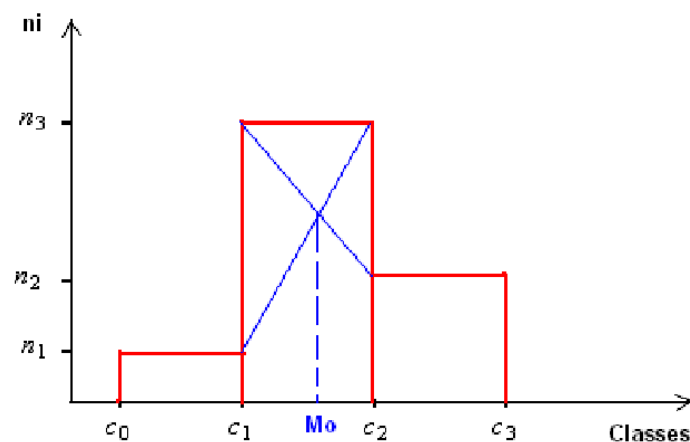
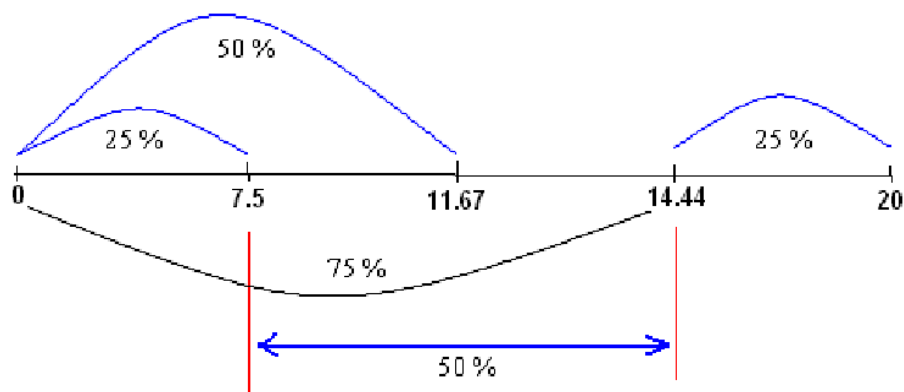
On obtient donc la répartition suivant :

L'intervalle interquartile est donc  $14.44 - 7.5 = 6.94$ .

### 2.1.4 Le mode

Le mode d'une variable statistique discrète  $X$  est la valeur  $x_i$  qui correspond à la fréquence la plus élevée, notée  $Mo$ . Si  $X$  est une variable statistique continue, on parle de *classe modale*.

#### – Exemple



Une enquête sur l'âge des élèves candidat à l'examen du bac à donne les résultats suivants :

Ages (années)	16	17	18	19	20	Total
Effectifs $n_i$	100	400	200	180	120	1000

L'effectif le plus élevé est 400, il correspond donc au mode  $Mo = 17$ .

◇ **Remarque 2**

*Une variable statistique peut avoir plusieurs modes, dans le cas où l'effectif le plus élevé correspond à plusieurs observations.*

*On dit aussi qu'une variable ne possède pas de mode si tous les effectifs partiels sont identiques.*

*Dans le cas d'une variable continue, le mode ne peut pas être trouvé d'une façon exacte, par contre on peut calculer une valeur approchée une fois qu'on a la classe modale qui est la classe qui a l'effectif le plus élevé si les classes sont d'amplitudes variables on utilise les effectifs rectifié ou les quantités  $\frac{f_i}{a_i}$ .*

Supposons que la classe modale est la classe  $[c_1, c_2[$  comprise entre les classes  $[c_0, c_1[$  et  $[c_2, c_3[$  de même amplitudes.

Une valeur approchée du mode est donnée par la relation suivante :

$$Mo = c_1 + (c_2 - c_1) \times \frac{(n_3 - n_1)}{(n_3 - n_1) + (n_3 - n_2)}.$$

## – Exemple

Soit le tableau donnant le poids (en kg) de 100 individus.

Poids en kg	[56 – 58[	[58 – 60[	[60 – 62[	[62 – 64[	[64 – 66[	Total
$n_i$	5	36	40	15	4	100

Toutes les classes ont la même amplitude, donc la classe modale est [60 – 62[.

$$\Rightarrow Mo = 60 + (62 - 60) \times \frac{(40-36)}{(40-36)+(40-15)} = 60.27.$$

## 2.1.5 La moyenne arithmétique

La moyenne arithmétique d'une variable statistique est égale à la somme des valeurs prises par cette variable, divisée par le nombre d'observations.

On la note par  $\bar{X}$  ou  $\mu$  telle que :

$$\bar{X} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

## – Exemple

La moyenne arithmétique des nombres 11, 10, 5, 12, 13, 8, 7 est

$$\bar{X} = \frac{11+10+5+12+13+8+7}{7} \approx 9.43.$$

## ◇ Remarque 3

La moyenne arithmétique est dite simple, car chaque valeur correspond à un individu et par opposition à la moyenne arithmétique pondérée.

## 2.1.6 La moyenne arithmétique pondérée

Soit une variable statistique dont les différentes modalités :  $x_1, x_2, \dots, x_k$ , aux quelles correspondent respectivement les effectifs :  $n_1, n_2, \dots, n_k$ , avec  $\sum_{i=1}^k n_i = n$ .

La moyenne arithmétique pondérée d'une telle variable est :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k n_i \times x_i = \sum_{i=1}^k f_i \times x_i$$

## ◇ Remarque 4

Si  $X$  est une variable continue alors  $\bar{X} = \frac{1}{n} \sum_{i=1}^k c_i \times x_i$ , avec  $c_i =$  centre de classe.

## – Exemple

Calculons la moyenne arithmétique pondérée de la variable étudiée dans l'exemple

Poids en kg	[56 – 58[	[58 – 60[	[60 – 62[	[62 – 64[	[64 – 66[	Total
$n_i$	5	36	40	15	4	100
Centres $c_i$	57	59	61	63	65	/

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k c_i \times x_i = \frac{1}{100} [5 \times 57 + 36 \times 59 + \dots + 4 \times 65] = 60.54.$$

## ► Calcul de la moyenne par changement de variable :



Quand les valeurs de la variable sont très grande, alors pour faciliter les calculs on fait un changement de variable en introduisant une nouvelle variable en introduisant une nouvelle variable  $X'$  dite variable auxiliaire, définie dans le cas d'une variable discrète par :  $x'_i = x_i - x_0$  (ou  $x'_i = x_i + x_0$ ), où  $x_0$  est une constante à définir, généralement elle est proche de la tendance centrale de la variable, elle est estimée par exemple, à partir du mode ou de la médiane. La moyenne de la variable auxiliaire est

$$\bar{X}' = \frac{1}{n} \sum_{i=1}^k n_i \times x'_i = \frac{1}{n} \sum_{i=1}^k n_i \times (x_i - x_0) = \bar{X} - x_0 \implies \bar{X} = \bar{X}' + x_0.$$

Si on pose  $x'_i = x_i + x_0 \implies \bar{X} = \bar{X}' - x_0$ .

– **Exemple**

Soit la variable statistique  $X$  dont la distribution est donnée par le tableau suivant :

$x_i$	10000	15000	20000	25000	40000	<b>Total</b>
$n_i$	20	50	70	40	20	200

Introduisons la nouvelle variable  $x'$  telle que  $x'_i = x_i - 20000$ , dont la distribution est la suivant

$x'_i$	-10000	-5000	0	5000	20000	<b>Total</b>
$n_i$	20	50	70	40	20	200

$$\bar{X}' = \frac{1}{n} \sum_{i=1}^5 n_i \times x'_i = 750 \implies \bar{X} = \bar{X}' + 20000 = 20750.$$

Dans le cas d'une variable continue on effectue le changement de variable suivant :  $y'_i = \frac{y_i - x_0}{y_0}$ , où  $x_0$  et  $y_0$  sont deux constantes à choisir de telle manière à réduire le volume des calculs.

De là on obtient  $\bar{Y} = y_0 \bar{Y}' + x_0$ .

– **Exemple**

Les prix de vente de 120 pièces mécaniques vendues dans un magasin de pièces détachées sont donnés par le tableau suivant :

<b>Prix en DA</b>	[200, 300[	[300, 400[	[400, 500[	[500, 600[	[600, 700[	Total
$n_i$	25	35	25	20	15	120
Centres $x_i$	250	350	450	550	650	/
$y'_i = \frac{y_i - 400}{100}$	-1;5	-0.5	0.5	1.5	2.5	/

On a posé  $x_0 = 400$  et  $y_0 = 100$ .

$$\bar{Y}' = \frac{1}{n} \sum_{i=1}^5 n_i \times y'_i = 0.21 \implies \bar{Y} = y_0 \bar{Y}' + x_0 = 100 \times 0.21 + 400 = 421.$$

### 2.1.7 La moyenne géométrique

La moyenne géométrique simple  $G$  d'une série de  $n$  valeurs  $x_1, x_2, \dots, x_n$  est la racine  $n^{\text{ième}}$  du produit de ces valeurs :

$$G = \sqrt[n]{x_1 \times x_2 \times \dots \times x_n}$$

En prenant le logarithme des deux membres on obtient :

$$\log G = \frac{1}{n} \sum_{i=1}^n \log x_i$$

Si on a une variable statistique ayant  $k$  valeurs  $x_1, x_2, \dots, x_k$  aux quelles correspondent respectivement les effectifs  $n_1, n_2, \dots, n_k$ , alors la moyenne géométrique pondérée de ces valeurs est :

$$G = \sqrt[n]{x_1^{n_1} \times x_2^{n_2} \times \dots \times x_k^{n_k}}$$

Ou encore

$$\log G = \frac{1}{n} \sum_{i=1}^k n_i \times \log x_i$$

– **Exemple**

La moyenne géométrique des valeurs 2, 3, 6, 11, 17 est

$$G = \sqrt[5]{2 \times 3 \times 6 \times 11 \times 17} = 6.12.$$

### 2.1.8 La moyenne harmonique

La moyenne harmonique simple  $H$  d'une série de  $n$  valeurs  $x_1, x_2, \dots, x_n$  est l'inverse de de la moyenne arithmétique des inverses des valeurs :

$$H = \frac{1}{\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Si on a une variable statistique ayant  $k$  valeurs  $x_1, x_2, \dots, x_k$  aux quelles correspondent respectivement les effectifs  $n_1, n_2, \dots, n_k$ , alors la moyenne harmonique pondérée de ces valeurs est :

$$H = \frac{1}{\frac{1}{n} \sum_{i=1}^k \frac{n_i}{x_i}} = \frac{n}{\sum_{i=1}^k \frac{n_i}{x_i}}$$

– **Exemple**

La moyenne harmonique des valeurs 1, 4, 8, 10, 12 est  $H = \frac{1}{\frac{1}{1} + \frac{1}{4} + \frac{1}{8} + \frac{1}{10} + \frac{1}{12}} = 3.2$ .

### 2.1.9 La moyenne quadratique

La moyenne quadratique simple  $MQ$  d'une série de  $n$  valeurs  $x_1, x_2, \dots, x_n$  est définie par :

$$MQ = \sqrt{x^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$$

Si on a une variable statistique ayant  $k$  valeurs  $x_1, x_2, \dots, x_k$  aux quelles correspondent respectivement les effectifs  $n_1, n_2, \dots, n_k$ , alors la moyenne quadratique pondérée de ces valeurs est :

$$MQ = \sqrt{x^2} = \sqrt{\frac{1}{n} \sum_{i=1}^n n_i \times x_i^2}$$

– **Exemple**

La moyenne quadratique des valeurs 7, 8, 10, 11, 14 est

$$MQ = \sqrt{\frac{1}{5} [(7)^2 + (8)^2 + (10)^2 + (11)^2 + (14)^2]} = 10.29.$$

◇ **Remarque 5**

On a la relation suivante entre ces différentes moyennes.

$$H \leq G \leq \bar{X} \leq MQ$$

## 2.2 Les caractéristiques de dispersion

Les caractéristiques de dispersion indiquent la disposition des observations autour de la moyenne.

Les caractéristiques de dispersion les plus utilisés sont : *L'étendu*, *la variance*, *l'écart type*, *l'écart absolue moyen* et *l'indice de concentration*

### 2.2.1 L'étendu

L'étendu d'un ensemble de valeurs est la différence entre la plus grande et la plus petite de ces valeurs  $E = x_{\max} - x_{\min}$ .

– **Exemple**

Soit la série statistique suivante : 3, 6, 9, 10, 14, 15, 6, 7, 11, 2, 5, 9.

L'étendu de cette série d'observations  $E = 15 - 2 = 13$ .

### 2.2.2 La variance

La variance d'une distribution statistique  $X$ , noté  $V(X)$  ou  $\sigma_X^2$  est la moyenne des carrés des écarts entre les observations  $x_i$  et la moyenne  $\bar{X}$  :

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

En développant cette expression on obtient une formule plus simple :

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

– **Exemple**

Soit la variable statistique dont la distribution est donnée par le tableau suivant :

$x_i$	1	2	3	4	5	<b>Total</b>
$n_i$	10	25	40	15	10	100

Calculons d'abord la moyenne  $\bar{X}$  :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k n_i x_i = \frac{1}{100} [(10 \times 1) + \dots + (10 \times 5)] = 2.9.$$

La variance  $V(X)$  est donc égale à :

$$V(X) = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - \bar{X}^2 = \frac{1}{100} [(10 \times 1^2) + \dots + (10 \times 5^2)] - (2.9)^2 = 1.19.$$

◇ **Remarque 6**

► En pratique il est préférable d'utiliser la formule développée de la variance, car elle rend les calculs plus faciles.

► Dans le cas d'une variable continue  $V(X) = \frac{1}{n} \sum_{i=1}^k n_i c_i^2 - \bar{X}^2$ , où  $c_i$  = centre de la classe  $i$ .

### 2.2.3 L'écart type

L'écart type  $\sigma_X$  d'une distribution statistique est égale à la racine carrée de la

$$\text{variance : } \sigma_X = \sqrt{V(X)} = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i (x_i - \bar{X})^2}$$

L'écart type est appelé par fois écart quadratique, car il représente la moyenne quadratique des écarts à la moyenne arithmétique.

### 2.2.4 L'écart absolue moyen

Soit la variable statistique  $X$ , ayant  $k$  modalités  $x_1, x_2, \dots, x_k$ , aux quelles correspondent respectivement les effectifs  $n_1, n_2, \dots, n_k$ .

« L'écart absolu moyen » de la variable statistique  $X$  est la moyenne arithmétique des valeurs absolues des écarts à la moyenne arithmétique :

$$e_{abs} = \frac{1}{n} \sum_{i=1}^k n_i \times |x_i - \bar{X}|$$

– **Exemple**

On donne la série des observations  $x_i$  : 5, 12, 13, 14, 2, 7, 18, 4, 10, 9. La moyenne arithmétique de cette variable statistique est

$$\bar{X} = \frac{1}{10} [5 + 12 + \dots + 9] = 9.4.$$

L'écart absolu moyen est donc égale à :

$$e_{abs} = \frac{1}{10} [|5 - 9.4| + |12 - 9.4| + \dots + |9 - 9.4|] = 4.$$

### 2.2.5 Le coefficient de variation

L'écart type, et la moyenne, s'expriment dans la même unité que la variable statistique, mais dans certain cas on peut être ramené à comparer les dispersions de distributions qui ne sont pas exprimées dans la même unité.

Pour cela, on utilise le coefficient de variation qui est défini par le rapport de l'écart type à la moyenne :  $CV = \frac{\sigma_X}{\bar{X}}$ .

– **Exemple**

La distribution des salaires observés dans deux entreprises ; une du sud et l'autre du nord sont donnés par les tableaux suivants :

Distribution des salaires dans l'usine du sud :

Salaires mensuels (DA)	[20000, 30000[	[30000, 35000[	[35000, 40000[	[40000, 50000[	Total
Effectifs $n_i$	30	20	40	10	100
Centres $c_i$	25000	32500	37500	45000	/

$$\bar{X}_1 = \frac{1}{n} \sum_{i=1}^k c_i \times x_i = 35000.$$

$$V(X_1) = \frac{1}{n} \sum_{i=1}^k n_i c_i^2 - \bar{X}_1^2 = 141500000 \implies \sigma_{X_1} = 11895.38.$$

Distribution des salaires dans l'usine du nord :

Salaires mensuels (DA)	[10000, 15000[	[15000, 20000[	[20000, 25000[	[25000, 40000[	Total
Effectifs $n_i$	40	30	20	10	100
Centres $c_i$	12500	17500	22500	32500	/

$$\bar{X}_2 = \frac{1}{n} \sum_{i=1}^k c_i \times x_i = 18000.$$

$$V(X_2) = \frac{1}{n} \sum_{i=1}^k n_i c_i^2 - \bar{X}_2^2 = 37250000 \implies \sigma_{X_2} = 6103.28.$$

Les coefficients des variations sont :

$$CV_1 = \frac{\sigma_{X_1}}{\bar{X}_1} = \frac{11895.38}{35000} = 0.355. \quad CV_2 = \frac{\sigma_{X_2}}{\bar{X}_2} = \frac{6103.28}{18000} = 0.339.$$

La dispersion relatives des salaires est presque la même dans les deux entreprises..

# Calculs des probabilités

---

## 3.1 Analyse combinatoire

Dans ce chapitre, on développe quelques techniques permettant de déterminer sans dénombrement direct le nombre de résultats possibles d'une expérience particulière, ou encore le nombre d'éléments d'un ensemble particulier. De telles techniques reçoivent souvent le nom d'analyse combinatoire.

### 3.1.1 Principe fondamental de l'analyse combinatoire

Si une procédure quelconque peut être représentée de  $n_1$  façons différentes, si après cette procédure, une seconde procédure peut être représentée de  $n_2$  façons différentes, et si ensuite une troisième procédure peut être représentée de  $n_3$  façons différentes, et ainsi de suite, alors le nombre de façons différentes permettant d'exécuter les procédures dans l'ordre indiqué est égal au produit  $n_1 \times n_2 \times n_3 \times \dots$

**Exemple :**

Supposons qu'une plaque d'immatriculation contient deux lettres distinctes suivies de trois chiffres dont le premier est différent de zéro. Combien de plaques différentes peut-on imprimer ?

Il y a 26 façons différentes d'imprimer la première lettre, 25 façons différentes d'imprimer la seconde lettre (puisque la première lettre imprimée ne peut plus être choisie pour la deuxième), 9 façons différentes d'imprimer le premier chiffre et dix façons différentes d'imprimer les deux autres chiffres. On en déduit que l'on peut imprimer  $26 \times 25 \times 9 \times 10 \times 10 = 585000$  plaques différentes.

### 3.1.2 Arrangements

Soit un ensemble de  $n$  éléments tout distincts  $\{a, b, \dots, s\}$ .

**Définition 1 :** On appelle « arrangement avec répétition » de  $p$  éléments choisis parmi  $n$ , toute disposition ordonnée de  $p$  éléments parmi les  $n$  avec répétition d'un ou de plusieurs éléments.

Le nombre d'arrangements de  $p$  éléments parmi  $n$ , noté  $\tilde{A}_n^p$  est égal à :

$$\tilde{A}_n^p = n^p = \underbrace{n \times n \times \dots \times n}_{p \text{ fois}}$$

### Exemples

1. Un numéro d'appel est composé de 8 chiffres, ce numéro est une disposition ordonnée avec répétition de 8 éléments parmi 10  $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$  donc un numéro est un arrangement avec répétition.

Le nombre de numéros possible est  $\tilde{A}_{10}^8 = 10^8$ .

2. Combien de mots différents de 4 lettres au plus, qu'on peut former avec un alphabet de (02) lettres ?

Un mot de  $p$  lettres choisi parmi deux est un arrangement avec répétition de  $p$  parmi  $n = 2 \{a, b\}$ .

► Nombre de mots d'une (01) lettre =  $\tilde{A}_2^1 = 2 \rightarrow (a), (b)$ .

► Nombre de mots de deux (02) lettres =  $\tilde{A}_2^2 = 2^2 = 4 \rightarrow (a, a), (a, b), (b, a), (b, b)$ .

► Nombre de mots de trois (03) lettres =  $\tilde{A}_2^3 = 2^3 = 8$ .

► Nombre de mots de quatre (04) lettres =  $\tilde{A}_2^4 = 2^4 = 16$ .

Donc au total il y a  $2 + 4 + 8 + 16 = 30$  mots formés d'au plus 4 lettres parmi 2.

**Remarque :** le nombre d'éléments  $p$  d'un arrangement avec répétition peut être  $\geq n$ , mais ce n'est pas le cas d'un arrangement sans répétition.

**Définition 2 :** On appelle « arrangement sans répétition » (ou simplement arrangement) de  $p$  éléments choisis parmi  $n$ , une disposition ordonnée sans répétition de  $p$  éléments parmi les  $n$ .

### Remarques :

► Dans un arrangement sans répétition les  $p$  éléments sont tous distincts,  $1 \leq p \leq n$ .

► L'ensemble des arrangements sans répétition est inclus dans l'ensemble des arrangements avec répétition.

Si  $A_n^p$  désigne le nombre d'arrangement sans répétition de  $p$  éléments parmi  $n$  alors

$$A_n^p \leq \tilde{A}_n^p.$$

$$\begin{aligned} A_n^p &= \frac{n!}{(n-p)!}; \quad 1 \leq p \leq n \\ &= n(n-1) \dots (n-(p-1)) \end{aligned}$$

### Exemples :

1. Quel est le nombre de tiercés dans l'ordre, d'une course de 10 chevaux.

L'ordre est important et on ne peut pas avoir de répétition, donc il s'agit d'arrangement sans répétition de trois éléments choisis parmi 10, leur nombre est égale à  $A_{10}^3 = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \times 9 \times 8 = 720$  tiercés possible.

2. Trois joueurs lancent un dé à 6 faces.

– Combien, y a-t-il de résultats possible ?

– Combien, y a-t-il de résultats où les points obtenus sont tous distincts ?

Le nombre de résultats possible est le nombre d'arrangements avec répétition de 3 parmi 6,  $\tilde{A}_6^3 = 6^3 = 216$ .

Le nombre de résultats différents est le nombre d'arrangements sans répétition de 3 parmi 6,  $A_6^3 = \frac{6!}{3!} = 120$ .

### 3.1.3 Permutations

Soit un ensemble de  $n$  éléments distincts  $\{a, b, \dots, s\}$ .

**Définition 1 :** On appelle « permutation sans répétition » ( ou simplement permutation ) de ces  $n$  éléments toute disposition ordonnée de l'ensemble de ces  $n$  éléments.

Chacun de ces éléments figure dans un ordre déterminé, une et une seule fois.

Donc une permutation est un arrangement sans répétition de  $n$  éléments choisis parmi  $n$ .

Le nombre de permutation possible de  $n$  éléments est

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

Avec convention  $0! = 1$ .

**Exemple :**

Trouver le nombre de configuration possible à l'arrivée d'une course de 10 chevaux.

Un arrivée est une permutation de 10 chevaux  $\implies$  Le nombre d'arrivée possible est donc  $P_{10} = 10! = 3628800$ .

Soit un ensemble de  $n$  éléments formés de  $k$  groupes différents, d'éléments identique

$(\underbrace{a, a, \dots, a}_{n_1}), (\underbrace{b, b, \dots, b}_{n_2}), \dots, (\underbrace{s, s, \dots, s}_{n_k})$  tels que  $n_1 + n_2 + \dots + n_k = n$ .

**Définition 2 :** On appelle « permutation avec répétition » de ces  $n$  éléments une disposition ordonnée de l'ensemble de ces  $n$  éléments et dans chaque permutation avec répétition la lettre  $a$  figure  $n_1$  fois, la lettre  $b$  figure  $n_2, \dots$ , la lettre  $s$  figure  $n_k$  fois.

Le nombre de ces permutations avec répétition est noté  $P_n(n_1, n_2, \dots, n_k)$ .

Cas où  $k = 2$

On a deux groupes  $(\underbrace{a, a, \dots, a}_{n_1}), (\underbrace{b, b, \dots, b}_{n_2}), n_1 + n_2 = n$

Soit par exemple  $n = 4, n_1 = n_2 = 2$  donc  $(a, b, b, a)$  est une permutation avec répétition.

Pour obtenir le nombre de permutations possible, il faut diviser le nombre de permutations sans répétition des  $n$  éléments, soit  $n!$ , par le nombre des permutations sans répétition des  $n_1$  éléments  $a$  entre eux.

Puis par le nombre des permutations sans répétition des  $n_2$  éléments  $b$  entre eux, ....

Par le nombre des permutations sans répétition des  $n_k$  éléments  $s$  entre eux. D'où :

$$P_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!}$$

**Exemples :**

1. Trouver le nombre des permutations du nombre 11222.

$$11222 = \left\{ (\underbrace{1, 1}_{n_1=2}), (\underbrace{2, 2, 2}_{n_2=3}) \right\} \implies P_5(2, 3) = \frac{5!}{2!3!} = 10$$

11222 12122 12212 12221 21122

21212 21221 22112 22121 22211

2. On jet 12 dés et on appelle « résultat » une suite ordonnée de 12 points obtenus. Combien y a-t-il de résultats au total et parmi eux combien y a-t-il de résultats où chaque face est obtenue 2 fois.

Un résultat est un arrangement avec répétition des 12 éléments choisis parmi 6, leur nombre est  $\tilde{A}_{12}^6 = 6^{12}$ .



Un résultat où chaque face est obtenue deux fois est une permutation des 12 éléments scindés en 6 groupes d'effectifs  $n_1 = n_2 = \dots = n_6 = 2$  leur nombre est  $P_{12}(2, 2, 2, 2, 2, 2) = \frac{12!}{2!2!2!2!2!2!}$  résultats.

### 3.1.4 Combinaisons

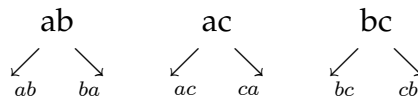
Soit un ensemble de  $n$  éléments distincts  $\{a, b, \dots, s\}$ .

**Définition 1 :** On appelle « combinaison sans répétition » ( ou plus simplement combinaison ) de  $p$  éléments choisis parmi les  $n$ , une disposition non ordonnée sans répétition de  $p$  éléments choisis parmi  $n$ , avec  $1 \leq p \leq n$ .

#### Exemple 1

$n = 3, \{a, b, c\}, p = 2$ .

Il y a 3 combinaisons de 2 parmi 3



On remarque qu'une combinaison engendre  $p!$  arrangements sans répétition, il y a donc  $p!$  fois plus d'arrangements que de combinaisons.

Donc, si  $C_n^p$  désigne le nombre de combinaisons de  $p$  parmi  $n$  alors  $p!C_n^p = A_n^p = \frac{n!}{(n-p)!}$ , d'où

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

#### Remarque :

- ▶  $C_n^0 = C_n^n = 1$
- ▶  $C_n^p = C_n^{n-p}$
- ▶  $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$

#### Exemple 2

Trouver le nombre de tiercés dans le désordre, dans une course de 10 chevaux.

Un tiercé dans le désordre est une combinaison de 3 éléments choisis parmi 10 leur nombre est  $C_{10}^3 = \frac{10!}{3!(10-3)!} = 120$ .

**Définition 2 :** On appelle « combinaison avec répétition » de  $p$  éléments choisis parmi les  $n$ , une disposition non ordonnée avec répétition éventuellement, de  $p$  éléments choisis parmi les  $n$  éléments.

#### Exemple 3

$n = 3, \{a, b, c\}, p = 2$ .

On a 6 combinaisons avec répétition :  $(a, a), (b, b), (c, c), (a, b), (a, c), (b, c)$ .

#### Remarque

$p$  peut être supérieur à  $n$  alors que, dans le cas des combinaisons sans répétition,  $p$  est au plus égal à  $n$ .

Le nombre des combinaisons avec répétition de  $p$  éléments choisis parmi  $n$  est donné la formule suivante :  $\tilde{C}_n^p = C_{n+p-1}^p$ .

## 3.2 Calculs des probabilités

*Le calcul des probabilité est la base essentielle pour toute approche de la théorie des probabilités et des statistiques. C'est une suite naturelle de la statistique descriptive dans laquelle on introduit de façon empirique les notions intuitives du langage de probabilité. La théorie des probabilité est considérée comme une science dont l'objet consiste à étudier des phénomènes qui dépendent du hasard.*

### 3.2.1 Espace probabilisable

#### Expérience aléatoire

Une expérience est qualifié d'aléatoire si on ne peut pas prévoir à l'avance son résultat et répétée dans des conditions identiques, elle peut donner lieu à des résultats différents.

#### Remarque

*Le résultat de l'épreuve dépend d'un ensemble de conditions dans les quelles s'effectue cette épreuve. Ces conditions doivent être pratiquement identiques, lorsqu'on répète l'expérience afin de s'assurer de la compatibilité des résultats.*

L'ensemble de tous les résultats possibles d'une épreuve est dit espace fondamental ou espace des événement (appelé aussi référentiel) noté  $\Omega$ .

#### Exemple 1

- ▶ Le jet d'un dé peut donner l'ensemble fondamental suivant :  $\Omega = \{1, 2, 3, 4, 5, 6\}$
- ▶ Une pièce de monnaie :  $\Omega = \{pile, face\}$ .

#### Événements élémentaires et composés

- ▶ Un événement élémentaire est un sous ensemble de  $\Omega$ , réduit à un seul élément.
- ▶ Un événement composé est un sous ensemble d'événements élémentaires.

#### Exemple 2

Pour le jet d'un dé. Avoir le chiffre 6 est événement élémentaire.

Avoir un chiffre pair est un événement composé de trois événements élémentaires : 2, 4 et 6.

#### Réalisation d'un événement

Considérons une expérience aléatoire d'ensemble fondamental  $\Omega$ . Soit  $A$  un événement quelconque associé à cette épreuve et considérons un résultat noté  $\omega$ . Lorsque le résultat  $\omega \in A$ , on dira que l'événement  $A$  s'est réalisé. Dans le cas contraire ( $\omega \notin A$ ), on dira que l'événement  $A$  ne s'est pas réalisé.

#### Événement incompatible (disjoint)

Deux événement  $A$  et  $B$  sont dits « incompatible » ou « disjoint » si la réalisation de l'un exclue de l'autre, c'est-à-dire  $A \cap B = \phi$ .

#### Exemple 3

On lance un dé et soient les événements :

$A$  « avoir un chiffre pair » =  $\{2, 4, 6\}$

$B$  « avoir un chiffre impair » =  $\{1, 3, 5\}$

$A \cap B = \phi$ , donc  $A$  et  $B$  sont incompatibles.

### Système complet d'événement

Soient  $A_1, A_2, \dots, A_n$  une famille d'événements. On dit que cette famille forme « un système complet » d'événements si :

$$- A_i \cap A_j = \phi, \quad \forall i \neq j,$$

$$- \bigcup_{i=1}^n A_i = \Omega.$$

### Algèbre des événements

On appelle « algèbre » ou « tribu » sur un espace fondamental  $\Omega$ , qu'on notera  $\mathcal{A}$ , une famille d'événements vérifiant :

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

$$- \text{Si } A \in \mathcal{A} \text{ alors } \bar{A} \in \mathcal{A}, \text{ où } \bar{A} \text{ est l'événement contraire de } A.$$

$$- \text{Si } A_i \in \mathcal{A}, \quad i \in I \text{ fini, alors } \bigcup_{i \in I} A_i \in \mathcal{A}.$$

#### Propriétés

$$- \phi \in \mathcal{A}.$$

$$- \text{Si } A_i \in \mathcal{A}, \quad i \in I \text{ fini, alors } \bigcap_{i \in I} A_i \in \mathcal{A}.$$

### Exemples 4

► La famille  $\{\phi, \Omega\}$  est une tribu sur  $\Omega$ . (c'est la plus petite tribu sur  $\Omega$ ).

► La famille  $\{\phi, A, \bar{A}, \Omega\}$  est une tribu sur  $\Omega$ , appelée tribu de Bernoulli.

► La famille des parties de  $\Omega$  est une tribu sur  $\Omega$ . (c'est la plus grande tribu sur  $\Omega$ ).

### Espace probabilisable

Un couple  $(\Omega, \mathcal{A})$ , constitué d'un ensemble fondamental  $\Omega$  et une tribu  $\mathcal{A}$  est appelé « espace probabilisable ». Les éléments de  $\mathcal{A}$  sont appelés « événements » et ceux de  $\Omega$  « événements élémentaires ».

### concept de probabilité

Définir une probabilité c'est associer à chaque événement  $A$ , un poids  $P(A)$  représentant « la chance qu'il a de se réaliser ».

#### a) Les différentes approches

il existe trois approches conceptuelles différentes des calculs de probabilités.

► *L'approche classique à priori*

Consiste à déterminer parmi tous les résultats possibles  $N$ , le nombre de résultats favorables  $n$  à la réalisation de l'événement  $A$  souhaité :

$$P(A) = \frac{n}{N} = \frac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}}$$

### Exemple 5

On jette une fois une pièce de monnaie, la probabilité de tomber sur face est :

$$P(\text{face}) = \frac{1}{2}.$$

Cette approche ne nécessitant pas une expérience est parfois désignée d'approche à priori.

► *L'approche fréquentiste ou à posteriori*

Cette approche se base sur les données fournies par un certain nombre d'observations ou d'expériences antérieures.

### Exemple 6

A priori la probabilité qu'une femme enceinte attende un garçon est  $\frac{1}{2}$ .

Or on constate, sur un million de naissances, qu'il y a 515000 garçons et 485000 filles.

La probabilité à posteriori (après l'expérience) d'avoir un garçon est égale à 0.515.

### b) Définition théorique

on appelle « probabilité » sur l'espace probabilisable  $(\Omega, \mathcal{A})$ , toute application  $P$  définie de  $(\Omega, \mathcal{A})$ , dans l'intervalle  $[0, 1]$  et vérifiant ;

–  $P(\Omega) = 1$ .

– Pour tout ensemble fini d'événements incompatibles  $A_1, A_2, \dots, A_n$  on a :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad (\text{Propriété d'additivité}).$$

En définissant une probabilité  $P$  sur  $(\Omega, \mathcal{A})$ , on obtient un espace dit « probabilisé » ou « espace de probabilité »  $(\Omega, \mathcal{A}, P)$ .

► Conséquence de l'axiome d'additivité

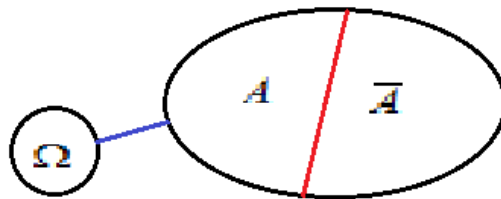
Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité, et  $A, B$  et  $C$  trois événements quelconques de  $\mathcal{A}$

– Probabilité de l'événement contraire :

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

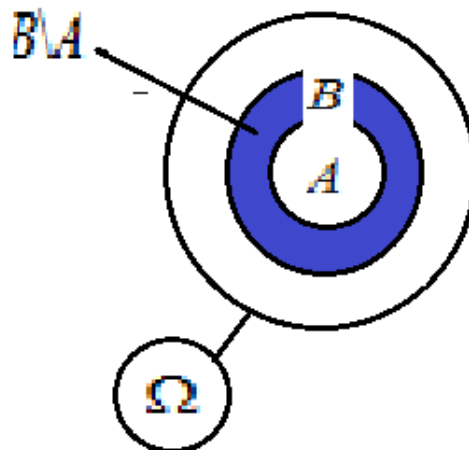
En effet, on a :  $A \cup \bar{A} = \Omega$  et  $A \cap \bar{A} = \phi$ , donc  $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$ .

– Probabilité de deux événements liés par l'inclusion :



Si  $A \subset B$  alors  $P(A) \leq P(B)$ .

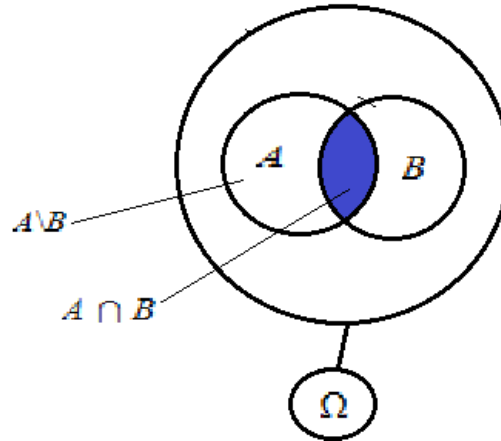
Nous avons  $B = A \cup (B \setminus A)$  et  $A \cap (B \setminus A) = \phi$ , donc  $P(B) = P(A) + P(B \setminus A)$  avec  $P(B \setminus A) \geq 0$ .



– Probabilité de différence de deux événements :

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

On a :  $A = (A \setminus B) \cup (A \cap B)$  et  $(A \setminus B) \cap (A \cap B) = \phi$

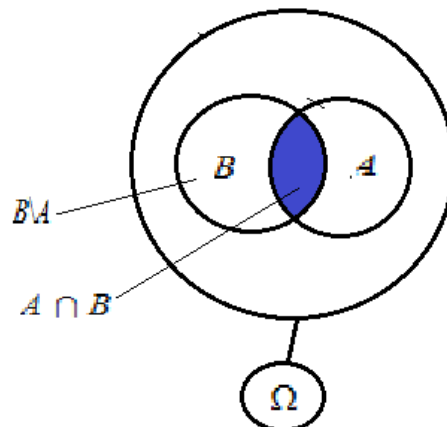


– Probabilité de la réunion de deux événements :

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

Car  $A \cup B = A \cup (B \setminus A)$  et

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \setminus A) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$



### 3.2.2 Lien entre la théorie des probabilités et des ensembles

Les résultats de la théorie des ensembles peuvent être interprétés sous forme d'événements en théorie des probabilités.

Ensembles	Événements
$A$	L'événement $A$ est réalisé
$\bar{A}$	L'événement $A$ n'est pas réalisé
$A \cup B$	L'événement $A$ ou l'événement $B$ est réalisé
$A \cap B$	L'événement $A$ et l'événement $B$ sont réalisés
$\phi$	L'événement impossible
$\Omega$	L'événement est certain
$A \cap B = \phi$	$A$ et $B$ sont deux événements incompatibles (mutuellement exclusif) leur réalisation simultanée est impossible.
$A \Delta B$	Un des événements seulement se réalise.

### 3.2.3 Construction d'une probabilité

– Les événements élémentaires  $\{\omega\}_{\omega \in \Omega}$  sont dits équiprobables s'ils ont tous la même probabilité, soit  $k$ .

On suppose bien sûr que  $\Omega$  est fini de cardinal  $n$ , muni de la tribu  $\mathcal{P}(\Omega)$ , on a alors :

$$1 = P\left(\bigcup_{i=1}^n \omega_i\right) = \sum_{i=1}^n P(\omega_i) = nk \implies k = \frac{1}{n}.$$

Si un événement  $A \in \mathcal{P}(\Omega)$  est constitué de  $n_A$  éléments alors :

$$\begin{aligned} P(A) &= P\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} P(\{\omega\}) = \frac{n_A}{n} = \frac{\text{Card}(A)}{\text{Card}(\Omega)} \\ &= \frac{\text{Nombre de cas favorable à la réalisation de } A}{\text{Nombre de cas possible}} \\ &= \frac{|A|}{|\Omega|}. \end{aligned}$$

#### Exemple 6

Soit  $(\Omega, \mathcal{P}(\Omega), P)$  un espace probabilisé et supposons que les événements élémentaires  $\{\omega_i\}$  sont équiprobables de probabilité  $P$ ,  $\text{Card}(\Omega) = n$ ,

$$\begin{aligned} P(\Omega) &= P\left(\bigcup_{i=1}^n \omega_i\right) = \sum_{i=1}^n P(\omega_i) = nP = 1 \implies P = \frac{1}{n} \\ P(A) &= P(\{\omega_1, \omega_2, \omega_3\}) = P(\omega_1) + P(\omega_2) + P(\omega_3) = \frac{3}{n}. \end{aligned}$$

#### Remarque

Cette probabilité  $P$  s'appelle probabilité uniforme sur  $(\Omega, \mathcal{P}(\Omega))$ .

### 3.2.4 Probabilités conditionnelles, indépendance et probabilités composées

#### Probabilités conditionnelles

Etant donné l'espace probabilisé  $(\Omega, \mathcal{A}, P)$  et un événement  $B$  de probabilité non nulle, on appelle : « probabilité conditionnelle » de  $A$  sachant  $B$ , soit  $P(A/B)$ , le rapport  $P(A/B) = \frac{P(A \cap B)}{P(B) \neq 0}$ .

L'application  $A \rightarrow P(A/B)$  défini sur  $(\Omega, \mathcal{A})$  est une probabilité.

#### Indépendance

##### Définition :

On dit que l'événement  $A$  est indépendant de l'événement  $B$  si :  $P(A/B) = P(A)$ , on a alors  $P(A \cap B) = P(A) \cdot P(B) \implies P(B/A) = P(B)$ .

En d'autre terme  $B$  est indépendant de  $A$ .

Ainsi on peut écrire,  $A$  et  $B$  sont indépendant 2 à 2  $\iff P(A \cap B) = P(A) \cdot P(B)$ .

#### Indépendance mutuelle

##### Définition :

Soient  $A_1, A_2, \dots, A_n$  des éléments de  $\mathcal{A}$ , ils sont dits « mutuellement indépendant » si pour toute partie  $I$  de l'ensemble  $\{1, 2, \dots, n\}$ , on a :

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

#### Exemple 7

On lance 2 dés, soient

1) L'événement  $A$  : « Avoir (6, 6) », et l'événement  $B$  : « Avoir un couple de nombre pairs ».

$$P(A) = \frac{1}{36}, P(B) = \frac{9}{36}.$$

On lance 2 dés, on apprend que  $B$  s'est réalisé, on s'intéresse alors à l'événement  $A$  sachant que  $B$  est réalisé, que l'on note  $A/B$ .

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/36}{9/36} = \frac{1}{9}.$$

Car  $P(A \cap B) = P(\{6, 6\}) = \frac{1}{36}$ .

2) Soient les événements :  $A$  : « Le 1<sup>er</sup> dé obtient un point impair »,  $B$  : « Le 2<sup>ème</sup> dé obtient un point impair »,  $C$  : « La somme des points est impair ».

Donc on a  $P(A) = \frac{1}{36}$ ,  $P(B) = \frac{18}{36}$ ,  $P(C) = \frac{18}{36} = \frac{1}{2}$ .

$$P(A \cap B) = \frac{9}{36} = \frac{1}{4} = P(A) \cdot P(B)$$

$$P(A \cap C) = \frac{9}{36} = \frac{1}{4} = P(A) \cdot P(C)$$

$$P(B \cap C) = \frac{9}{36} = \frac{1}{4} = P(B) \cdot P(C)$$

$A, B$  et  $C$  sont 2 à 2 indépendants, mais il ne sont pas mutuellement indépendants, car  $P(A \cap B \cap C) = 0$ .

**Probabilités composés****Proposition :**

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

Résultat que l'on généralise à  $n$  événements  $A_1, A_2, \dots, A_n$  :

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

**Formule de Bayes**

► 1<sup>er</sup> formule de Bayes :

$$P(B/A) = \frac{P(A/B) P(B)}{P(A)} \quad (1)$$

**Définition :**

$A_1, A_2, \dots, A_n$  forment un « système complet d'événements » si  $A_1, A_2, \dots, A_n$  forment une partition de  $\Omega$ , c'est-à-dire :

$$\begin{cases} A_i \cap A_j = \phi, & i \neq j \\ \bigcup_{i=1}^n A_i = \Omega. \end{cases}$$

► 2<sup>ème</sup> formule de Bayes :

Soit  $(B_i)_{1 \leq i \leq n}$  un système complet d'événements et  $A$  un événement de probabilité non nulle, de la formule (1) on déduit la 2<sup>ème</sup> formule de Bayes dite aussi (« probabilités des causes »)

$$P(B_i/A) = \frac{P(A/B_i) P(B_i)}{\sum_{i=1}^n P(A/B_i) P(B_i)}, \quad 1 \leq i \leq n \quad (2)$$



## Série de TD 3

**Ex 1 :** Dans les élections des comités pédagogiques, 5 étudiants délégués et 6 étudiants comme adjoints se présentent.

- Combien de comités comprenant délégués et adjoints qu'on peut former ?

**Sol**

$A_1$  : l'ensemble des étudiants se présentant pour délégué.

$A_2$  : l'ensemble des étudiants se présentant pour adjoint.

On a  $|A_1| = 5$ ;  $|A_2| = 6$

L'ordre est important et on ne peut pas avoir des répétitions, le nombre des comités possible est  $A_5^1 \times A_6^1 = 5 \times 6 = 30$ .

**Ex 2 :** Quelle est la capacité ( théorique ) d'un réseau téléphonique sachant qu'un numéro est constitué de 6 chiffres.

**Sol**

la capacité ( théorique ) d'un réseau téléphonique à 6 chiffres est un arrangement avec répétition de 6 éléments choisis parmi 10  $\{0, 1, \dots, 9\}$ ,  $\tilde{A}_{10}^6 = 10^6$ .

**Ex 3 :** Trois joueurs A, B, C lancent chacun un dé à 6 faces.

1) Combien y a-t-il de résultats possibles ?

2) Combien y a-t-il de résultats où les points obtenus sont tous différents ?

**Sol**

1) 3 joueurs A, B, C lancent chacun un dé à 6 faces, un résultat possible est de la forme  $(a, b, c)$  où :

$a$  : est le résultat obtenu par le joueur A.

$b$  : est le résultat obtenu par le joueur B.

$c$  : est le résultat obtenu par le joueur C.

c'est donc un arrangement avec répétition de 3 éléments choisis parmi 6.

Donc le nombre de résultats possible est le nombre d'arrangements avec répétition  $\tilde{A}_6^3 = 6^3$ .

2) Si les points obtenus sont différents, le nombre de résultats est le nombre d'arrangements sans répétition de 3 chiffres parmi 6, donc  $A_6^3 = \frac{6!}{3!} = 120$ .

**Ex 4 :** Quel est le nombre d'ordres différents dont un médecin peut recevoir 7 personnes arrivées au même temps.

**Sol**

Le nombre d'ordres est le nombre de permutations de 7 personnes :  $P_7 = 7!$ .

**Ex 5 :** Une urne contient 9 boules numérotées de 1 à 9, on prélève 4 boules l'une après l'autre sans remise.

Quel est le nombre de chiffres qu'on peut former avec les boules prélevées ?

**Sol**

On tire 4 boules l'une après l'autre, donc l'ordre est important, le tirage est sans remise, donc on peut pas avoir des répétitions  $\implies$  il s'agit d'arrangements sans répétitions de 4 éléments choisis parmi 9. Leurs nombres est :

$$A_9^4 = \frac{9!}{(9-4)!} = 9 \times 8 \times 7 \times 6 = 3024.$$

**Ex 6 :** Une assemblée veut élire un bureau de 3 membres, il y a 5 candidats.

1) Combien de bureaux peut on former ?

2) Si parmi les 5 candidats il y a 2 femmes, combien de bureaux comportant exactement 2 femmes peuvent on former ?

**Sol**

1) une disposition est formée de 3 membres, l'ordre n'est pas important et on ne peut pas avoir des répétitions, donc il s'agit de combinaison sans répétitions,  $n = 5, p = 3$ .

Le nombre de bureaux possible est  $C_5^3 = \frac{5!}{3!(5-3)!} = 20$ .

2) Le nombre de bureaux comportant exactement 2 femmes est égale à  $C_2^2 \cdot C_3^1 = 3$ .

**Ex 7 :** On dispose d'un lot de pièces dont trois sont bonnes et 3 sont défectueuses.

1) Combien peut on réaliser d'échantillons différents de 3 pièces ?

2) Combien parmi ces échantillons, ceux qui contiennent 3 pièces bonnes ?

3) Combien d'échantillon qui contiennent au moins une pièce bonne ?

**Sol**

6 pièces  $\left\{ \begin{array}{l} 3 \text{ bonnes} \\ 3 \text{ défectueuses} \end{array} \right.$

1) L'ordre n'est pas important et on ne peut pas avoir des répétitions.

Il s'agit donc de combinaisons sans répétitions.

Le nombre d'échantillons différents de 3 pièces est  $C_6^3$ .

2) Le nombre d'échantillons contenant 3 pièces bonnes est  $C_3^3 \cdot C_3^0$ .

3) Le nombre d'échantillons contenant au moins une pièce bonne est

$\{(1B, 2D) \text{ ou } (2B, 1D) \text{ ou } (3B, 0D)\}$

$C_3^1 \cdot C_3^2 + C_3^2 \cdot C_3^1 + C_3^3 \cdot C_3^0$ .

**Ex 8 :** Un hôpital veut engager 4 médecins dans 4 spécialités différentes. Six médecins se présentent.

Combien de choix s'offrent au responsable de l'embauche dans les cas suivants :

1) Les six médecins sont polyvalents (pouvant occuper un des 4 postes).

2) Un seul est polyvalent pour les 4 postes, Les 5 autres le sont seulement dans 3 branches, les même pour les cinq.

**Sol**

L'ordre n'est pas important et on ne peut pas avoir des répétitions.

1)  $C_6^3$ , 2)  $C_1^1 \cdot C_4^3$ .

**Ex 9 :** On dispose de  $n_1$  cubes blancs et  $n_2$  cubes rouges.

On les range «au hasard» dans des cases numérotées de 1 à  $n_1 + n_2$ .

1) Quel est le nombre  $N$  des dispositions possibles distinctes? Rép :  $N = \frac{(n_1+n_2)!}{n_1! \cdot n_2!}$

2) Quel est le nombre de dispositions lorsque  $k$  cubes blancs déterminés se trouvent à des places fixées. Rép :  $N' = \frac{(n_1+n_2-k)!}{(n_1-k)! \cdot n_2!}$ .

**Ex 10 :** Les 12 tomes d'une encyclopédie sont placés au hasard.

1) Combien y a-t-il de manières de les classer.

2) Parmi ces classements, combien y en a-t-il où les tomes 1 et 2 se trouvent côté à côté.

**Sol**

1) L'ordre est important et on ne peut pas avoir des répétitions, de plus  $n = p$  donc il s'agit de permutation sans répétition de 12 éléments, leur nombre est  $P_{12} = 12! \quad 2) 2! \cdot 10!$ .

# TD et EXAMENS

---

## 4.1 Série de TD 1

### Série de TD 1

#### Ex 1

Pour chacune des variables suivantes préciser si elle est :

►Qualitative ►Quantitative discrète ►Quantitative continue.

1. Revenu annuel. 2. Lieu de résidence. 3. Citoyenneté. 4. Âge. 5. Sexe. 6. Pointure en chaussures. 7. Couleur des yeux. 8. état matrimonial. 9. Tour de taille. 10. Nombre de langues parlées.

#### Sol

Ces Variables peuvent être classés selon leurs natures en 3 catégories :

Qualitative	Quantitative discrète	Quantitative continue
2. Lieu de résidence	6. Pointure en chaussures	1. Revenu annuel
3. Citoyenneté	10. Nombre de langues parlées	4. Âge
5. Sexe		9. Tour de taille
7. Couleur des yeux		
8. état matrimonial		

#### Ex 2

Une enquête menée sur la répartition des groupes sanguins a donné les résultats suivants :

Groupe sanguin	A	B	AB	O
Effectifs	219	123	78	242

- 1) Quels sont les effectifs partiels ?
- 2) Quel est l'effectif total ?
- 3) Quelles sont les valeurs du caractère ?
- 4) De quel type est-il ?
- 5) Tracer avec précision le diagramme correspondant à ce caractère.

#### Sol

1) Les effectifs partiels sont les effectifs correspondant à chacune des modalités :

$$n_1 = 219, n_2 = 123, n_3 = 78, n_4 = 242.$$

2) L'effectif total est la somme des effectifs partiels :

$$n = n_1 + n_2 + n_3 + n_4 = 662.$$

3) Les valeurs du caractère représentent les différentes modalités : A, B, AB et O.

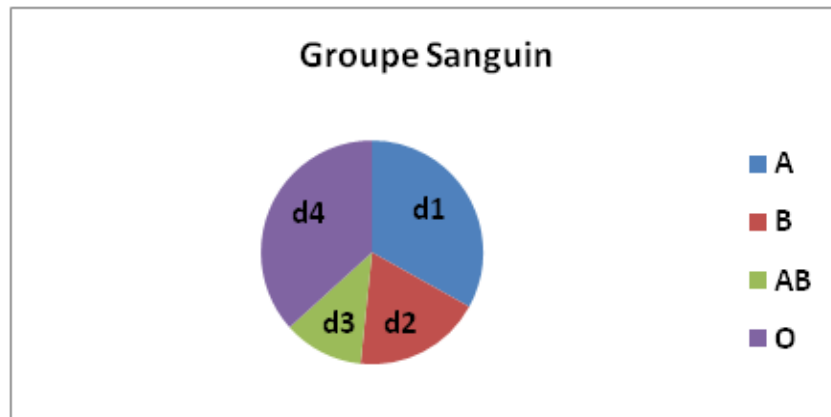
4) Le caractère étudié est qualitatif.

5) Ce caractère peut être représenté graphiquement par le diagramme circulaire.

Calculons les degrés des secteurs correspondant aux différentes modalités :

$$d_i = \frac{n_i}{n} \times 360^\circ, \quad i = 1, 2, 3, 4$$

$$d_1 \simeq 119, d_2 \simeq 67, d_3 \simeq 42, d_4 \simeq 132$$



### Exercice 3

On a relevé les nombres des points obtenus par 50 candidats dans un concours de recrutement, les résultats sont les suivants :

59 101 76 86 99 87 99 101 77 58 87 66 82

59 77 81 89 114 97 87 99 77 77 99 79 88

107 86 89 76 71 80 83 79 100 98 77 84 81

89 85 86 85 83 85 67 83 88 75 100

1) Quel est le type du caractère à l'étude ici ?

2) Regrouper les valeurs de cet échantillon en 5 classes, de telle sorte que 30 soit la borne inférieure de la première classe et que l'amplitude de chacune des classes soit de 20, bâtir ensuite la table des effectifs de cette distribution.

3) Tracer l'Histogramme des effectifs.

## 4.2 Série de TD 3 (Analyse Combinatoire)

### Série de TD 3

**Ex 1 :** Dans les élections des comités pédagogiques, 5 étudiants délégués et 6 étudiants comme adjoints se présentent.

- Combien de comités comprenant délégués et adjoints qu'on peut former ?

**Sol**

$A_1$  : l'ensemble des étudiants se présentant pour délégué.

$A_2$  : l'ensemble des étudiants se présentant pour adjoint.

On a  $|A_1| = 5$ ;  $|A_2| = 6$

L'ordre est important et on ne peut pas avoir des répétitions, le nombre des comités possible est  $A_5^1 \times A_6^1 = 5 \times 6 = 30$ .

**Ex 2 :** Quelle est la capacité ( théorique ) d'un réseau téléphonique sachant qu'un numéro est constitué de 6 chiffres.

**Sol**

la capacité ( théorique ) d'un réseau téléphonique à 6 chiffres est un arrangement avec répétition de 6 éléments choisis parmi 10  $\{0, 1, \dots, 9\}$ ,  $\tilde{A}_{10}^6 = 10^6$ .

**Ex 3 :** Trois joueurs A, B, C lancent chacun un dé à 6 faces.

1) Combien y a-t-il de résultats possibles ?

2) Combien y a-t-il de résultats où les points obtenus sont tous différents ?

**Sol**

1) 3 joueurs A, B, C lancent chacun un dé à 6 faces, un résultat possible est de la forme  $(a, b, c)$  où :

$a$  : est le résultat obtenu par le joueur A.

$b$  : est le résultat obtenu par le joueur B.

$c$  : est le résultat obtenu par le joueur C.

c'est donc un arrangement avec répétition de 3 éléments choisis parmi 6.

Donc le nombre de résultats possible est le nombre d'arrangements avec répétition  $\tilde{A}_6^3 = 6^3$ .

2) Si les points obtenus sont différents, le nombre de résultats est le nombre d'arrangements sans répétition de 3 chiffres parmi 6, donc  $A_6^3 = \frac{6!}{3!} = 120$ .

**Ex 4 :** Quel est le nombre d'ordres différents dont un médecin peut recevoir 7 personnes arrivées au même temps.

**Sol**

Le nombre d'ordres est le nombre de permutations de 7 personnes :  $P_7 = 7!$ .

**Ex 5 :** Une urne contient 9 boules numérotées de 1 à 9, on prélève 4 boules l'une après l'autre sans remise.

Quel est le nombre de chiffres qu'on peut former avec les boules prélevées ?

**Sol**

On tire 4 boules l'une après l'autre, donc l'ordre est important, le tirage est sans remise, donc on peut pas avoir des répétitions  $\implies$  il s'agit d'arrangements sans répétitions de 4 éléments choisis parmi 9. Leurs nombres est :

$$A_9^4 = \frac{9!}{(9-4)!} = 9 \times 8 \times 7 \times 6 = 3024.$$

**Ex 6 :** Une assemblée veut élire un bureau de 3 membres, il y a 5 candidats.

1) Combien de bureaux peut on former ?

2) Si parmi les 5 candidats il y a 2 femmes, combien de bureaux comportant exactement 2 femmes peuvent on former ?

**Sol**

1) une disposition est formée de 3 membres, l'ordre n'est pas important et on ne peut pas avoir des répétitions, donc il s'agit de combinaison sans répétitions,  $n = 5, p = 3$ . Le nombre de bureaux possible est  $C_5^3 = \frac{5!}{3!(5-3)!} = 20$ .

2) Le nombre de bureaux comportant exactement 2 femmes est égale à  $C_2^2 \cdot C_3^1 = 3$ .

**Ex 7 :** On dispose d'un lot de pièces dont trois sont bonnes et 3 sont défectueuses.

1) Combien peut on réaliser d'échantillons différents de 3 pièces ?

2) Combien parmi ces échantillons, ceux qui contiennent 3 pièces bonnes ?

3) Combien d'échantillon qui contiennent au moins une pièce bonne ?

**Sol**

6 pièces  $\left\{ \begin{array}{l} 3 \text{ bonnes} \\ 3 \text{ défectueuses} \end{array} \right.$

1) L'ordre n'est pas important et on ne peut pas avoir des répétitions. Il s'agit donc de combinaisons sans répétitions.

Le nombre d'échantillons différents de 3 pièces est  $C_6^3$ .

2) Le nombre d'échantillons contenant 3 pièces bonnes est  $C_3^3 \cdot C_3^0$ .

3) Le nombre d'échantillons contenant au moins une pièce bonne est

$\{(1B, 2D) \text{ ou } (2B, 1D) \text{ ou } (3B, 0D)\}$

$C_3^1 \cdot C_3^2 + C_3^2 \cdot C_3^1 + C_3^3 \cdot C_3^0$ .

**Ex 8 :** Un hôpital veut engager 4 médecins dans 4 spécialités différentes. Six médecins se présentent.

Combien de choix s'offrent au responsable de l'embauche dans les cas suivants :

1) Les six médecins sont polyvalents (pouvant occuper un des 4 postes).

2) Un seul est polyvalent pour les 4 postes, Les 5 autres le sont seulement dans 3 branches, les même pour les cinq.

**Sol**

L'ordre n'est pas important et on ne peut pas avoir des répétitions.

1)  $C_6^3$ , 2)  $C_1^1 \cdot C_4^3$ .

**Ex 9 :** On dispose de  $n_1$  cubes blancs et  $n_2$  cubes rouges.

On les range «au hasard» dans des cases numérotées de 1 à  $n_1 + n_2$ .

1) Quel est le nombre  $N$  des dispositions possibles distinctes ? Rép :  $N = \frac{(n_1+n_2)!}{n_1! \cdot n_2!}$

2) Quel est le nombre de dispositions lorsque  $k$  cubes blancs déterminés se trouvent à des places fixées. Rép :  $N' = \frac{(n_1+n_2-k)!}{(n_1-k)! \cdot n_2!}$ .

**Ex 10 :** Les 12 tomes d'une encyclopédie sont placés au hasard.

1) Combien y a-t-il de manières de les classer.

2) Parmi ces classements, combien y en a-t-il où les tomes 1 et 2 se trouvent côté à côté.

**Sol**

1) L'ordre est important et on ne peut pas avoir des répétitions, de plus  $n = p$  donc il s'agit de permutation sans répétition de 12 éléments, leur nombre est  $P_{12} = 12! \quad 2) 2! \cdot 10!$ .

## 4.3 Série de TD 4 (Calcul des probabilités)

### Série de TD 4

#### Exercice 1

Parmi les fonctions suivantes les quelles définissent une probabilité sur  $\Omega = \{a, b, c\}$ ?

- 1)  $P(a) = \frac{1}{4}, P(b) = \frac{1}{3}, P(c) = \frac{1}{2}$ ;
- 2)  $P(a) = \frac{1}{6}, P(b) = \frac{1}{3}, P(c) = \frac{1}{2}$ ;
- 3)  $P(a) = \frac{2}{3}, P(b) = -\frac{1}{3}, P(c) = \frac{2}{3}$ ;
- 4)  $P(a) = 0, P(b) = \frac{1}{3}, P(c) = \frac{2}{3}$ .

#### Sol

- 1)  $\frac{1}{4} + \frac{1}{3} + \frac{1}{2} = \frac{13}{12} \neq 1$  Non
- 2)  $\frac{1}{6} + \frac{1}{3} + \frac{1}{2} = 1$  Oui
- 3)  $P(b) = -\frac{1}{3} < 0$  Non
- 4)  $0 + \frac{1}{3} + \frac{2}{3} = 1$  Oui

#### Exercice 2

Soient  $A$  et  $B$  deux événements tels que :

$$P(A) = \frac{3}{8}, P(B) = \frac{1}{2}, P(A \cap B) = \frac{1}{4}.$$

Calculer  $P(A \cup B), P(\bar{A}), P(\bar{B}), P(\bar{A} \cap \bar{B}), P(\bar{A} \cup \bar{B})$ .

#### Sol

- $P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{8} + \frac{1}{2} - \frac{1}{4} = \frac{5}{8}$ .
- $P(\bar{A}) = 1 - P(A) = 1 - \frac{3}{8} = \frac{5}{8}$ .
- $P(\bar{B}) = 1 - P(B) = 1 - \frac{1}{2} = \frac{1}{2}$ .
- $P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - \frac{5}{8} = \frac{3}{8}$ .
- $P(\bar{A} \cup \bar{B}) = P(\overline{A \cap B}) = 1 - P(A \cap B) = 1 - \frac{1}{4} = \frac{3}{4}$ .

#### Exercice 3

Une urne contient 10 boules dont 4 rouges et 6 noires, on prélève 3 boules de cette urne. En distinguant le cas d'un tirage simultané et de tirages successifs, calculer la probabilité pour que :

- 1) Les trois boules soient de même couleur.
- 2) Deux boules exactement sont de couleur rouge.
- 3) Au moins une des boules est de couleur rouge.

#### Sol

1<sup>er</sup> Cas : Tirage simultané

La disposition  $\{R, R, N\} = \{N, R, R\}$  donc l'ordre n'est pas important, comme les boules sont supposées

numérotés, donc on ne peut pas avoir des répétitions, il s'agit donc de combinaisons sans répétitions de 3 boules choisis parmi 10.

$$\Omega = \{\{R, R, N\}, \{N, N, N\}, \dots\}, |\Omega| = C_{10}^3 = 120.$$

1)  $A$  «Les trois boules soient de même couleur»

$$A = \{\{R, R, R\} \text{ ou } \{N, N, N\}\}, |A| = C_6^3 + C_4^3 = 20 + 4 = 24.$$

$$\implies P(A) = \frac{|A|}{|\Omega|} = \frac{24}{120} = \frac{1}{5}.$$

2)  $B$  «Deux boules exactement sont de couleur rouge»

$$B = \{\{R, R, N\}\}, |B| = C_4^2 \cdot C_6^1 = 36.$$

$$\implies P(B) = \frac{|B|}{|\Omega|} = \frac{36}{120} = \frac{3}{10}.$$

3)  $C$  «Au moins une des boules est de couleur rouge»

$$C = \{\{1R, 2N\} \text{ ou } \{2R, 1N\}\}, |C| = C_4^1 \cdot C_6^2 + C_4^2 \cdot C_6^1 + C_4^3 = 76.$$

$$\implies P(C) = \frac{|C|}{|\Omega|} = \frac{76}{120} = \frac{19}{30}.$$

2<sup>ème</sup> Cas : Tirages successifs avec remise

Dans ce cas l'ordre est important et on peut avoir des répétitions, donc il s'agit d'arrangements avec répétitions.

$$|\Omega| = \tilde{A}_{10}^3 = 10^3 = 1000.$$

1) A «Les trois boules soient de même couleur»

$$A = \{(3N) \text{ ou } (3R)\}, \quad |A| = \tilde{A}_4^3 + \tilde{A}_6^3 = 4^3 + 6^3 = 64 + 216 = 280.$$

$$\implies P(A) = \frac{|A|}{|\Omega|} = \frac{280}{1000} = \frac{7}{25}.$$

2) B «Deux boules exactement sont de couleur rouge»

$$B = \{\{R, R, N\} \text{ ou } \{R, N, R\} \text{ ou } \{N, R, R\}\}$$

$$|B| = 3 \cdot \tilde{A}_4^2 \cdot \tilde{A}_6^1 = 3 \cdot 16 \cdot 6 = 288.$$

$$\implies P(B) = \frac{|B|}{|\Omega|} = \frac{288}{1000} = \frac{36}{125}.$$

3) C «Au moins une des boules est de couleur rouge»

$$C = \{\{1R, 2N\} \text{ ou } \{2R, 1N\} \text{ ou } (3R)\}$$

$$|C| = 3 \cdot \tilde{A}_4^1 \cdot \tilde{A}_6^2 + 3 \cdot \tilde{A}_4^2 \cdot \tilde{A}_6^1 + \tilde{A}_4^3 = 784$$

Autrement :

$\bar{C}$  «Toute les boules sont noires»

$$|\bar{C}| = \tilde{A}_6^3 = 216$$

$$\implies P(C) = 1 - P(\bar{C}) = 1 - \frac{216}{1000} = \frac{784}{1000} = \frac{98}{125}.$$

3<sup>ème</sup> Cas : Tirages successifs sans remise

Il s'agit d'arrangements sans répétitions.

$$|\Omega| = A_{10}^3 = \frac{10!}{(10-3)!} = 10 \times 9 \times 8 = 720.$$

$$1) A = \{(3R) \text{ ou } (3N)\}, \quad |A| = A_4^3 + A_6^3 = 24 + 120 = 144 \implies P(A) = \frac{144}{720} = \frac{1}{5}.$$

$$2) B = \{\{R, R, N\} \text{ ou } \{R, N, R\} \text{ ou } \{N, R, R\}\}$$

$$|B| = 3 \cdot A_4^2 \cdot A_6^1 = 3 \cdot 12 \cdot 6 = 216 \implies P(B) = \frac{216}{720} = \frac{3}{10}.$$

$$3) C = \{\{1R, 2N\} \text{ ou } \{2R, 1N\} \text{ ou } (3R)\}$$

$$\bar{C} = \{(3N)\}, \quad |\bar{C}| = A_6^3 = 120 \implies P(C) = 1 - P(\bar{C}) = 1 - \frac{120}{720} = \frac{600}{720} = \frac{5}{6}.$$

#### Exercice 4

On jette simultanément une pièce de monnaie et un dé.

1) Décrire l'ensemble fondamentale.

2) Soit A l'événement «Obtenir pile» et B l'événement «Obtenir un nombre < 3»

– Décrire  $\bar{A}$  et  $\bar{B}$ .

– Donner la décomposition de A, B et  $A \cap B$ .

3) Calculer  $P(A)$ ,  $P(B)$ ,  $P(A \cap B)$ ,  $P(A/B)$ .

4) A et B sont ils indépendants ?

**Sol**

1)

$$\Omega =$$

$$\{(1, P), (2, P), (3, P), (4, P), (5, P), (6, P), (1, F), (2, F), (3, F), (4, F), (5, F), (6, F)\}$$

$$2) A = \{(1, P), (2, P), (3, P), (4, P), (5, P), (6, P)\}, \quad |A| = 6.$$

$$B = \{(1, P), (1, F), (2, P), (2, F)\}, \quad |B| = 4.$$

$$\bar{A} \text{ «Avoir face»} = \{(1, F), \dots, (6, F)\}, \quad |\bar{A}| = 6.$$

$$\bar{B} \text{ «Avoir un nombre } \geq 3 \text{»} = \Omega - B, \quad |\bar{B}| = 8.$$

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

$$3) P(A) = \frac{|A|}{|\Omega|} = \frac{6}{12} = \frac{1}{2}, \quad P(B) = \frac{|B|}{|\Omega|} = \frac{4}{12} = \frac{1}{3}, \quad P(A \cap B) = \frac{|A \cap B|}{|\Omega|} = \frac{2}{12} = \frac{1}{6}.$$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/3} = \frac{1}{2}.$$



$$4) P(A/B) \stackrel{?}{=} P(A)$$

On a  $\frac{1}{2} = \frac{1}{2} \implies A$  et  $B$  sont indépendants.

Autrement

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

$$\frac{1}{6} = \frac{1}{2} \times \frac{1}{3}$$

#### Exercice 5

Une urne contient 4 boules rouges, 6 boules blanches et 10 boules noires, on tire 3 boules à la fois.

Calculer les probabilités des événements suivants :

- 1) Toutes les boules tirées sont rouges.
- 2) Toutes les boules sont noires.
- 3) 2 boules sont rouges et 1 blanche.
- 4) Une boule de chaque couleur.
- 5) Au moins une boule noire.

**Sol**

$$U(4R, 6B, 10N)$$

$$1) A \ll \text{Avoir 3 boules rouges} \gg, P(A) = \frac{C_4^3}{C_{20}^3} = \frac{4}{1140}.$$

$$2) B \ll \text{Toutes les boules noires} \gg, P(B) = \frac{C_{10}^3}{C_{20}^3} = \frac{120}{1140}.$$

$$3) C \ll \text{Avoir 2R et 1B} \gg, P(C) = \frac{C_4^2 \times C_6^1}{C_{20}^3} = \frac{36}{1140}.$$

$$4) D \ll \text{Une boule de chaque couleur} \gg, P(D) = \frac{C_4^1 \times C_6^1 \times C_{10}^1}{C_{20}^3} = \frac{240}{1140}.$$

$$5) H \ll \text{Au moins une boule noire} \gg, P(H) = \frac{C_{10}^1 \times C_{10}^2 + C_{10}^2 \times C_{10}^1 + C_{10}^3 \times C_{10}^0}{C_{20}^3} = \frac{1040}{1140}.$$

#### Exercice 6

On cherche une personne qui, avec la probabilité  $p/8$ , se trouve dans un quelconque des 8 étages d'un immeuble. On a exploré en vain les 7 premiers étages.

Quelle est la probabilité  $f(p)$  que la personne se trouve au 8<sup>ème</sup> étage ?

**Sol**

Soit l'événement  $A_i$  «La personne se trouve au  $i^{\text{ème}}$  étage»,  $i = 1, 2, \dots, 8$ .

$$\text{On a } P(A_i) = \frac{p}{8}.$$

Soit l'événement  $B$  «La personne se trouve au 7 premiers étages»

$$B = \bigcup_{i=1}^7 A_i \implies P(B) = \sum_{i=1}^7 P(A_i) = \frac{7p}{8}.$$

Nous on cherche  $P(A_8/\bar{B}) = ?$

$$P(A_8/\bar{B}) = \frac{P(A_8 \cap \bar{B})}{P(\bar{B})} = \frac{P(A_8)}{1 - P(B)} = \frac{p/8}{1 - 7p/8} \implies f(p) = \frac{p}{8 - 7p}.$$

#### Exercice 7

- 1) Quelle est la probabilité qu'une famille de six enfants ait trois garçons ?
- 2) Une famille à deux enfants dont au moins un garçon. Quelle est la probabilité que cette famille ait deux garçons ?

**Sol**

1) En admettant que les sexes des enfants consécutifs sont indépendants et que chaque enfant à la probabilité  $\frac{1}{2}$  d'être un garçon.

$$\Omega = \{(G, F, F, F, G, F) \dots\}, |\Omega| = 2^6 = 64.$$

$$A : \ll \text{Avoir 3G} \gg, |\Omega| = C_6^3; \quad P(A) = \frac{C_6^3}{2^6} = \frac{20}{64} = \frac{5}{16}.$$

$$2) \Omega = \{(G, G), (G, F), (F, G)\}, |\Omega| = 3.$$

$$B : \ll (G, G) \gg, P(B) = \frac{1}{3}.$$

**Exercice 8**

Soit  $\Omega$  un ensemble fondamental,  $A$  et  $B$  deux événements tels que :

$$P(A) = 0.3; \quad P(B) = 0.2; \quad P(A \cap B) = 0.1$$

Déterminer les probabilités des événements suivants :

- 1) Au moins un des événements  $A$  et  $B$  se réalise.
- 2) Aucun des événements  $A, B$  ne se réalise.
- 3)  $A$  ne se réalise pas mais  $B$  se réalise.

**Sol**

$$1) P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.2 - 0.1 = 0.4$$

$$2) P(\overline{A \cap B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0.4 = 0.6$$

$$3) P(\overline{A} \cap B) = ?$$

$$\text{On a } \overline{A} \cap B = B - A \cap B \implies P(\overline{A} \cap B) = P(B) - P(A \cap B) = 0.2 - 0.1 = 0.1$$

## 4.4 Examen 2012

### Examen Statistique 2012

#### Exercice 1

On relevé les revenus annuels d'un échantillon de 20 personnes. Les valeurs sont les suivantes (en centaines de DA) :

360	380	280	340	360	440	260	340	320	360
300	360	300	360	340	360	380	400	320	380

I)

- a) Quelle est la nature de la variable ?  
 b) Calculer la moyenne arithmétique  $\bar{X}$  et la médiane  $Me$  de cette distribution.

II)

- a) Ranger ces résultats en classe de même amplitude de 40. (prenant la borne inférieure de la première classe = 26).  
 b) Tracer l'Histogramme de cette distribution et la courbe des fréquences cumulées.  
 c) Calculer le 1<sup>er</sup> quartile et 2<sup>ème</sup> quartile de cette distribution.  
 d) En posant :  $y_i = \frac{x_i - 260}{40}$ , calculer  $\bar{Y}$ ,  $V(Y)$  et  $\sigma_Y$ . En déduire  $\bar{X}$ ,  $V(X)$  et  $\sigma_X$ .  
 e) Expliquer pourquoi la moyenne et la médiane obtenues en II) sont différentes de celle obtenue en I).

Sol

I)

- a) le caractère étudié :  $X$  le revenu annuel, son type : quantitatif discret.

Revenu $\times 10^2$ DA	260	280	300	320	340	360	380	400	440	Total
Effectifs $n_i$	1	1	2	2	3	6	3	1	1	20

- b)\* la moyenne

$$\bar{X} = \frac{1}{n} \sum_{i=1}^9 n_i \times x_i = \frac{1}{20} \left[ \begin{array}{l} 260 \times 1 + 280 \times 1 + 300 \times 2 + 320 \times 2 + \\ 340 \times 3 + 360 \times 6 + 380 \times 3 + 400 \times 1 + 440 \times 1 \end{array} \right]$$

$$= \frac{6940}{20} \Rightarrow \bar{X} = 347. \text{ (34700 DA)}$$

\* La médiane : - On doit d'abord ordonner la série statistique dans l'ordre croissant :

Comme  $n = 20$  est pair alors  $n = 2p = 2 \times 10$

$$\Rightarrow p = 10 \Rightarrow Me = \frac{x_{10} + x_{11}}{2} = \frac{360 + 360}{2} \Rightarrow Me = 360. \text{ (36000 DA)}$$

II)

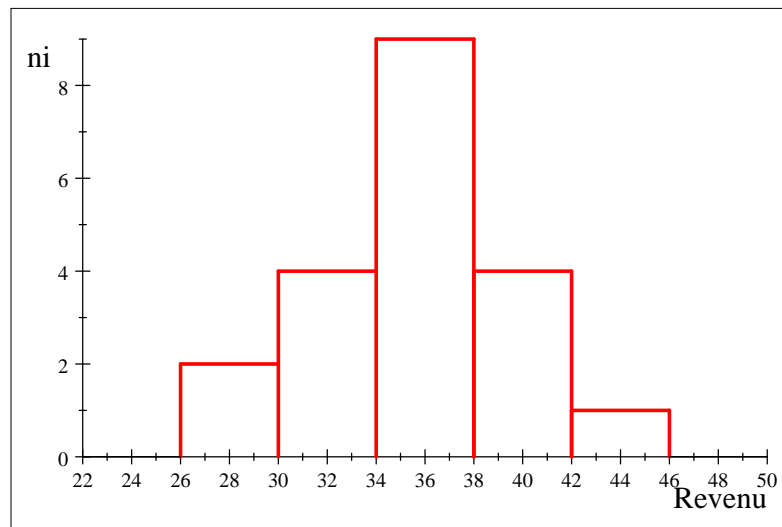
- a) Rangement des résultats en classe de même amplitude  $a_i = 40$  :

- le nombre de classes qu'on peut obtenir :  $k = \frac{x_{\max} - x_{\min}}{40} \approx 5$  classes.

- Dressons d'abord le tableau statistique en calculons les centres de classes  $c_i$ , les fréquences  $f_i$ , les amplitudes  $a_i$  et les fréquences cumulées (ou fonction de répartition  $F(x)$ ) :

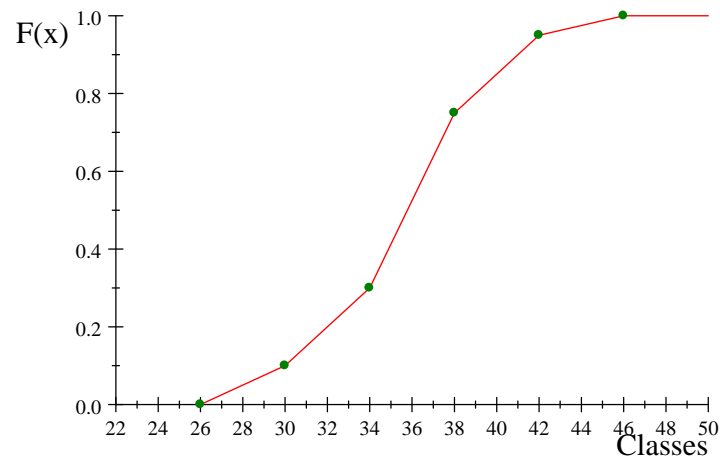
Revenu $\times 10^2$ DA	[260, 300[	[300, 340[	[340, 380[	[380, 420[	[420, 460[	Total
Effectifs $n_i$	2	4	9	4	1	20
Centres $c_i$	280	320	360	400	440	/
Fréquences $f_i$	0.1	0.2	0.45	0.2	0.05	1
F ; cum $F(x)$	0.1	0.3	0.75	0.95	1	/

- b)\* L'Histogramme : comme les classes ici sont de même amplitude, alors les hauteurs des rectangles de l'histogramme doivent être proportionnelles à la fréquence  $n_i$  :



Histogramme

\* La courbe des fréquences cumulées : on joint les points ayant pour abscisses la limite supérieure des classes et pour ordonnées les fréquences relatives cumulées correspondant à la classes considérée.



La courbe cumulative

c)\* Les quartiles : pour calculer les quartiles, on a besoin des fréquences cumulées  $F(x)$ .

- Le premier quartile  $Q_1$  : (localisation d'abord)

le premier quartile correspond à la valeur 0.25 de la fonction de répartition i.e

$$F(Q_1) = 0.25$$

$$\underbrace{F(300)}_{=0.1} \leq 0.25 \leq \underbrace{F(340)}_{=0.3} \implies Q_1 \in [300, 340[, \text{ alors}$$

$$Q_1 = 300 + (340 - 300) \frac{0.25 - 0.1}{0.3 - 0.1} = 330 \text{ (33000 DA)}$$

C'est-à-dire, 25% des personnes ont des revenus annuels inférieur à 33000 DA.

- Le deuxième quartile  $Me$  : (localisation d'abord)

la médiane correspond à la valeur 0.50 de la fonction de répartition i.e  $F(Me) = 0.50$

$$\underbrace{F(340)}_{=0.3} \leq 0.25 \leq \underbrace{F(380)}_{=0.75} \implies Me \in [340, 380[, \text{ alors}$$

$$Me = 340 + (380 - 340) \frac{0.5 - 0.3}{0.75 - 0.3} = 357.78 \quad (35778 \text{ DA})$$

C'est-à-dire, 50% des personnes ont des revenus annuels inférieur à 35778 DA.

d)

Revenu $\times 10^2$ DA	[260, 300[	[300, 340[	[340, 380[	[380, 420[	[420, 460[	Total
Fréquences $f_i$	0.1	0.2	0.45	0.2	0.05	1
$y_i = \frac{c_i - 260}{40}$	0.5	1.5	2.5	3.5	4.5	/

- Calculons :  $\bar{Y}$ ,  $V(Y)$  et  $\sigma_Y$

$$\bar{Y} = \sum_{i=1}^k f_i \times y_i = 0.1 \times 0.5 + 0.2 \times 1.5 + 0.45 \times 2.5 + 0.2 \times 3.5 + 0.05 \times 4.5 = 2.4$$

$$V(Y) = \sum_{i=1}^k f_i \times y_i^2 - (\bar{Y})^2 = \left[ \begin{array}{l} 0.1 \times (0.5)^2 + 0.2 \times (1.5)^2 + 0.45 \times (2.5)^2 + \\ 0.2 \times (3.5)^2 + 0.05 \times (4.5)^2 \end{array} \right] - (2.4)^2 =$$

$$6.75 - 5.76$$

$$\implies V(Y) = 0.99 \approx 1.$$

$$\text{D'où } \sigma_Y = \sqrt{V(Y)} \approx 1.$$

- Déduisons  $\bar{X}$ ,  $V(X)$  et  $\sigma_X$

$$\text{Comme } y_i = \frac{x_i - 260}{40} \implies \bar{Y} = \frac{\bar{X} - 260}{40} \implies \bar{X} = 40\bar{Y} + 260 \implies \bar{X} = 40(2.4) + 260 = 356.$$

$$V(X) = (40)^2 V(Y) \approx 1600. \text{ D'où } \sigma_X = \sqrt{V(X)} \approx 40.$$

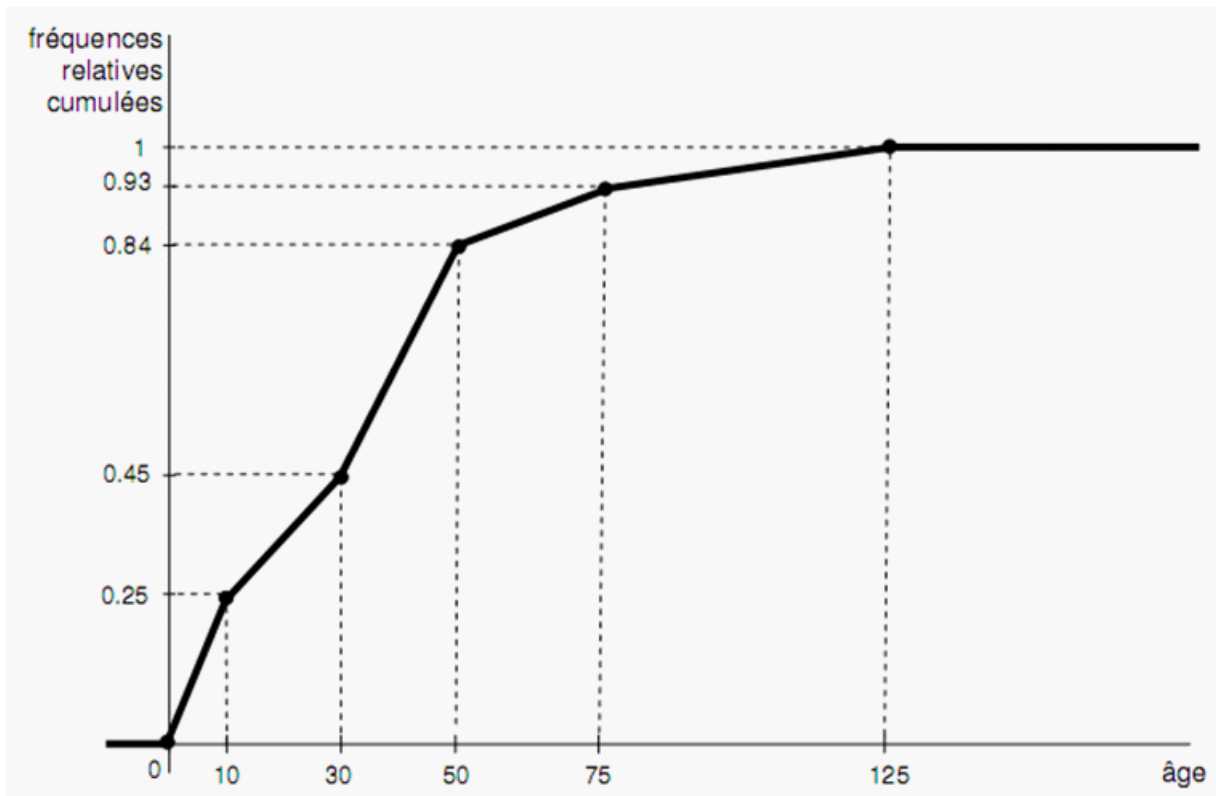
e) La différence tient au fait qu'en regroupant les valeurs en classes, on perd un peu d'information et aussi au fait qu'à l'intérieur des classes, les modalités ne sont pas équi-réparties.

### Exercice 2

Sur un échantillon d'habitants d'une certaine région, on a observé l'âge de ces

habitants. On a réparti ces individus en 5 classes d'âge (limites de classe :

0, 10, 30, 50, 75 et 125). On a alors pu obtenir le graphique suivant des fréquences relatives cumulées (« fonction de répartition ») :



- 1) Calculer la fréquence relative de chacune des cinq classes.
  - 2) Calculer la moyenne et la médiane de l'âge des individus de l'échantillon.
  - 3) Dessiner l'histogramme des fréquences.
- 1) Dressons d'abord le tableau statistique :

**Sol**

Classes	Fréquences cumulées $F(x)$	Amplitudes $a_i$	Centres $c_i$	Fréquences $f_i$	Fréquences rectifiés $f_i' = 10 \times \frac{f_i}{a_i}$
[0; 10[	0.25	10	5	0.25	0.25
[10; 30[	0.45	20	20	0.20	0.1
[30; 50[	0.84	20	40	0.39	0.195
[50; 75[	0.93	25	62.5	0.09	0.036
[75; 125[	1	50	100	0.07	0.014
<b>Total</b>	/	/	/	1	/

2) \* la moyenne :

$$\bar{X} = \sum_{i=1}^k f_i \times c_i = 0.25 \times 5 + 0.20 \times 20 + 0.39 \times 40 + 0.09 \times 62.5 + 0.07 \times 100 \implies \bar{X} = 33.475.$$

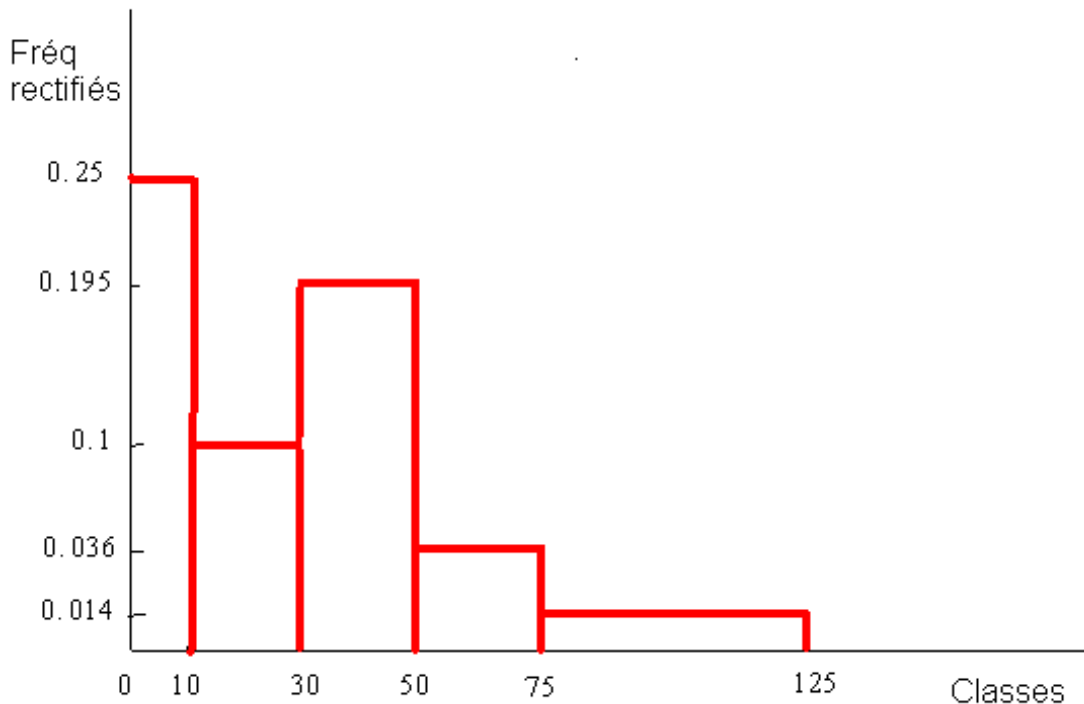
\* la médiane : On l'obtient en prenant l'abscisse du point d'ordonnée 0,50 sur la courbe des fréquences cumulées précédente.

- La valeur de la médiane s'obtient, dans la classe médiane [30; 50[, par interpolation linéaire entre les points (30; 0,45) et (50; 0,84) :

$$Me = 30 + (50 - 30) \frac{0.50 - 0.45}{0.84 - 0.45} = 32.56$$

3) L'histogramme des fréquences : Construction de l'histogramme : Comme les classes ici ne sont pas de même amplitude, alors, les hauteurs des rectangles de l'histogramme doivent être proportionnelles aux fréquences rectifiées

$$f_i^* = f_i \times \frac{\min(a_i)}{a_i} = 10 \times \frac{f_i}{a_i}$$



### Exercice 3 (Questions de cours)

Soit  $(\Omega, \mathcal{A}, P)$  un espace de probabilité, et  $A, B$  et  $C$  trois événements quelconques de  $\mathcal{A}$

Démontrer les axiomes de probabilité suivante

- 1)  $P(\bar{A}) = 1 - P(A)$ .
- 2) Si  $A \subset B$  alors  $P(A) \leq P(B)$ .
- 3)  $P(A \setminus B) = P(A) - P(A \cap B)$ .
- 4)  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ .

**Sol**

- 1) On a :  $A \cup \bar{A} = \Omega$  et  $A \cap \bar{A} = \phi$ , donc  $1 = P(\Omega) = P(A \cup \bar{A}) = P(A) + P(\bar{A})$ .
- 2) Nous avons  $B = A \cup (B \setminus A)$  et  $A \cap (B \setminus A) = \phi$ , donc  $P(B) = P(A) + P(B \setminus A)$  avec  $P(B \setminus A) \geq 0$ .
- 3) On a :  $A = (A \setminus B) \cup (A \cap B)$  et  $(A \setminus B) \cap (A \cap B) = \phi$ , donc  $P(A) = P(A \setminus B) + P(A \cap B)$ .
- 4) On a :  $A \cup B = A \cup (B \setminus A)$  et

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \setminus A) \\ &= P(A) + P(B) - P(A \cap B). \end{aligned}$$

## 4.5 Examen de Rattrapage 2012

### Rattrapage Statistique 2012

#### Exercice 1

La répartition en pourcentage des 50 employés d'une entreprise selon leurs salaires annuels (en 1000 DA) est donnée par le tableau suivant :

Salaires annuels $\times 10^3$ DA	Fréquences cumulées $F_i$ en %
[0, 30[	20
[30, 60[	48
[60, 90[	84
[90, 120[	100

- 1) Calculer les fréquences relatives et déduire les différents effectifs.
- 2) Tracer l'Histogramme de cette distribution.
- 3) Calculer le mode ( $Mo$ ).
- 4) Quel est le salaire médian ( $Me$ )? Interpréter le résultat.
- 5) Calculer les trois quartiles  $Q_1$ ,  $Q_2$  et  $Q_3$ .
- 6) Déterminer le salaire annuel moyen.
- 7) Calculer la variance et l'écart-type.
- 8) Utilisons la méthode d'interpolation linéaire, trouver le nombre des employés ayant un salaire inférieur à 45000 DA.

#### Sol

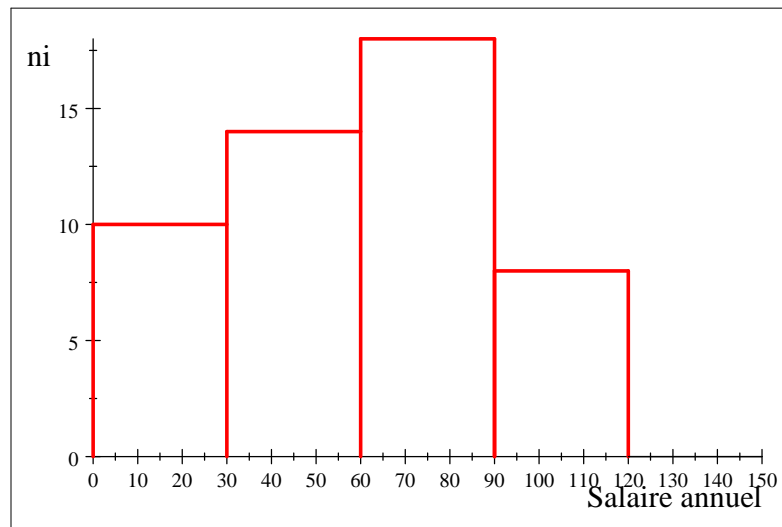
1)

$$f_i = F_i - F_{i-1}, \quad f_i = \frac{n_i}{N} \implies n_i = f_i \times N$$

$[e_{i-1}, e_i[$	$F_i$	$f_i$	$n_i$	$c_i$
[0, 30[	0.20	0.20	10	15
[30, 60[	0.48	0.28	14	45
[60, 90[	0.84	0.36	18	75
[90, 120[	1	0.16	8	105
Total	/	1	50	/

- 2) L'Histogramme : comme les classes ici sont de même amplitude, alors les hauteurs des rectangles de l'histogramme doivent être proportionnelles aux effectifs  $n_i$  (ou les fréquences  $f_i$ )





Histogramme

3) le mode ( $Mo$ )

$\max(n_i) = 18 \implies$  la classe modale est la classe  $[60, 90[$

Utilisons la formule de calcul

$$Mo = 60 + (90 - 60) \times \frac{18 - 14}{(18 - 14) + (18 - 8)}$$

$\implies Mo = 68.571$  (68571 DA).

4) le salaire médian ( $Me$ )

la médiane correspond à la valeur 0.50 de la fonction de répartition i.e  $F(Me) = 0.50$

$\underbrace{F(60)}_{=0.48} \leq 0.5 < \underbrace{F(90)}_{=0.84} \implies Me \in [60, 90[$ , alors

$$Me = 60 + (90 - 60) \frac{0.5 - 0.48}{0.84 - 0.48} = 61.667 \text{ (61667 DA)}$$

Le salaire médian moyen est 61667 DA.

Interprétation : Il y a 25 personnes qui ont un salaire annuel inférieur à 61667 DA.

5) les trois quartiles  $Q_1$ ,  $Q_2$  et  $Q_3$ .

- Le premier quartile  $Q_1$  : (localisation d'abord)

le premier quartile correspond à la valeur 0.25 de la fonction de répartition i.e

$F(Q_1) = 0.25$   
 $\underbrace{F(30)}_{=0.20} \leq 0.25 < \underbrace{F(60)}_{=0.48} \implies Q_1 \in [30, 60[$ , alors

$$Q_1 = 30 + (60 - 30) \frac{0.25 - 0.20}{0.48 - 0.20} = 35.357 \text{ (35357 DA)}$$

- Le deuxième quartile  $Q_2 = Me = 61.667$  (61667 DA)

- Le troisième quartile  $Q_3$  : (localisation d'abord)

le troisième quartile correspond à la valeur 0.75 de la fonction de répartition i.e

$F(Q_3) = 0.75$

$$\underbrace{F(60)}_{=0.48} \leq 0.75 < \underbrace{F(90)}_{=0.84} \implies Q_3 \in [60, 90[, \text{ alors}$$

$$Q_3 = 60 + (90 - 60) \frac{0.75 - 0.48}{0.84 - 0.48} = 82.5 \text{ (82500 DA)}$$

6) le salaire annuel moyen

$$\bar{X} = \sum_{i=1}^k f_i \times c_i = 0.20 \times 15 + 0.28 \times 45 + 0.36 \times 75 + 0.16 \times 105$$

$$\implies \bar{X} = 59.4 \text{ (59400 DA)}$$

7) la variance et l'écart-type

$$V(X) = \sum_{i=1}^k f_i \times c_i^2 - (\bar{X})^2 = \left[ \begin{array}{l} 0.20 \times (15)^2 + 0.28 \times (45)^2 + \\ 0.36 \times (75)^2 + 0.16 \times (105)^2 \end{array} \right] - (59.4)^2 =$$

$$4401 - 3528.4 = 872.6$$

$$\implies V(X) = 872.6.$$

$$\text{D'où } \sigma_X = \sqrt{V(X)} = \sqrt{872.6} = 29.540 \text{ (29540 DA)}.$$

8) le nombre des employés ayant un salaire inférieur à 45000 DA.

soit  $F(45) = x$ .

D'après la méthode d'interpolation linéaire

$$\frac{F(45) - F(30)}{45 - 30} = \frac{F(60) - F(30)}{60 - 30} \iff \frac{x - 0.20}{45 - 30} = \frac{0.48 - 0.20}{60 - 30}$$

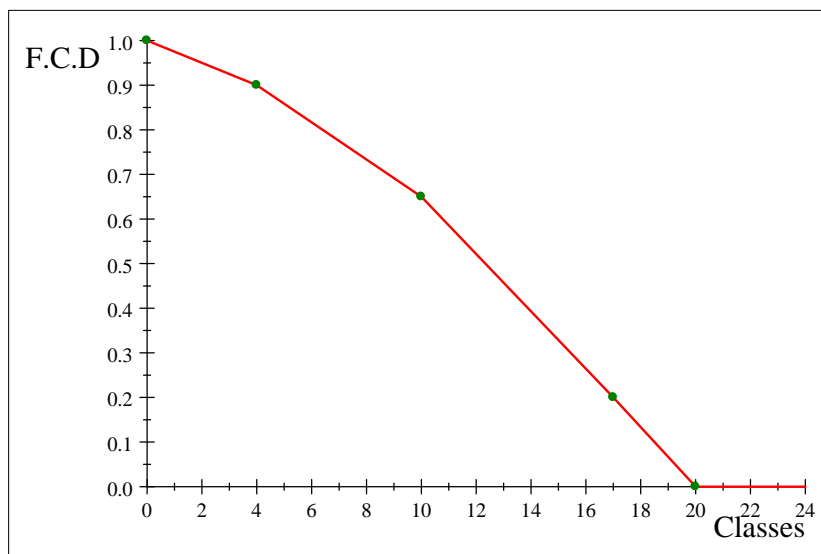
$$\iff x = 0.34 = 34\%$$

le nombre des employés ayant un salaire inférieur à 45000 DA égale à  $50 \times 0.34 = 17$  employés.

### Exercice 2

On a relevé les notes au module de Statistique, observées sur un échantillon d'étudiants du 1<sup>er</sup> année LMD MI.

On a réparti ces individus en 4 classes de note (limites de classe : 0, 4, 10, 17 et 20). On a alors pu obtenir le graphique suivant des fréquences relatives cumulées décroissantes :



Courbe cumulative décroissante

- 1) Calculer la fréquence relative de chacune des quatre classes.
- 2) Calculer la moyenne et la médiane de note des individus de l'échantillon.
- 3) Dessiner l'histogramme des fréquences.

**Sol**

1) Dressons d'abord le tableau statistique :

$$f_i = 1 - F_{i+1} \downarrow$$

Classes	F.C.D $F_i \downarrow$	F.C.C $F_i \uparrow$	Amplitudes $a_i$	Centres $c_i$	Fréquences $f_i$	Fréquences rectifiés $f_i^{\wedge} = 3 \times \frac{f_i}{a_i}$
[0; 4[	1	0.1	4	2	0.10	0.075
[4; 10[	0.9	0.35	6	7	0.25	0.125
[10; 17[	0.65	0.8	7	13.5	0.45	0.193
[17; 20[	0.2	1	3	18.5	0.20	0.20
<b>Total</b>	/	/	/	/	1	/

2) \* la moyenne :

$$\bar{X} = \sum_{i=1}^k f_i \times c_i = 0.10 \times 2 + 0.25 \times 7 + 0.45 \times 13.5 + 0.20 \times 18.5 \implies \bar{X} = 11.725.$$

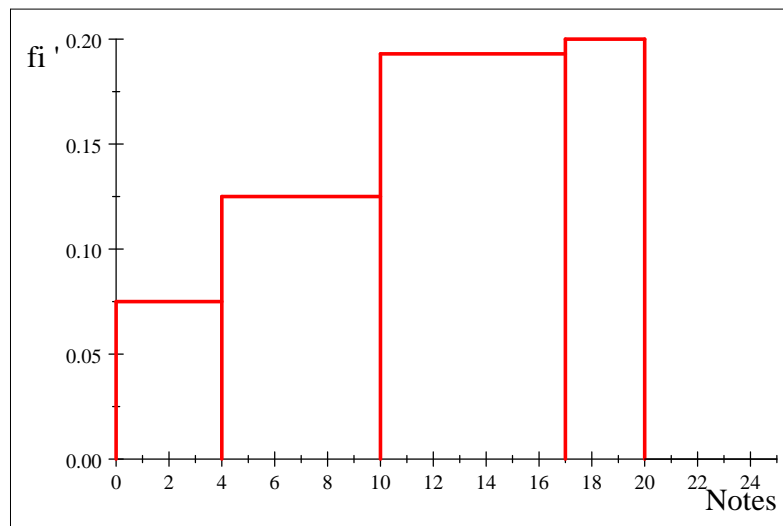
\* la médiane : On l'obtient en prenant l'abscisse du point d'ordonnée 0,50 sur la courbe des fréquences cumulées décroissantes précédente.

- La valeur de la médiane s'obtient, dans la classe médiane [10; 17[, par interpolation linéaire entre les points (10; 0,35) et (17; 0,65) :

$$Me = 10 + (17 - 10) \frac{0.50 - 0.35}{0.65 - 0.35} = 13.5$$

3) L'histogramme des fréquences : Construction de l'histogramme : Comme les classes ici ne sont pas de même amplitude, alors, les hauteurs des rectangles de l'histogramme doivent être proportionnelles aux fréquences rectifiées

$$f_i^{\wedge} = f_i \times \frac{\min(a_i)}{a_i} = 3 \times \frac{f_i}{a_i}$$



### Exercice 3

Soient  $A$  et  $B$  des événements tels que  $P(A) = \frac{1}{4}$ ,  $P(A \cup B) = \frac{1}{3}$  et  $P(B) = p$ .

(1) Calculer  $p$  en supposant que  $A$  et  $B$  sont incompatibles (i.e  $A \cap B = \phi$ ),

(2) Calculer  $p$  en supposant que  $A$  et  $B$  sont indépendants (i.e

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

(3) Calculer  $p$  en supposant que  $A$  est un sous-ensemble de  $B$  (i.e  $A \subset B$ ).

(4) Pour quel valeur de  $p$ , on a  $P(A/B) > 0$  ?

**Sol**

$$(1) A \cap B = \phi \implies P(A \cup B) = P(A) + P(B) \implies \frac{1}{3} = \frac{1}{4} + p \implies p = \frac{1}{12}.$$

$$(2) A \text{ et } B \text{ sont indépendants} \implies P(A \cup B) = P(A) + P(B) - P(A) \times P(B)$$

$$\implies \frac{1}{3} = \frac{1}{4} + p - \frac{1}{4} \times p \implies p = \frac{1}{9}.$$

(3)

$$A \subset B \implies A \cap B = A \implies P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(B) \implies p = \frac{1}{3}.$$

$$(4) P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) + P(B) - P(A \cup B)}{P(B)} = \frac{\frac{1}{4} + p - \frac{1}{3}}{p} = \frac{p - \frac{1}{12}}{p} > 0 \implies p > \frac{1}{12}.$$

## 4.6 Examen 2011

### Examen stat 2011

#### Ex 1

On a relevé les notes obtenues à l'examen de Statistique dans un groupe de 20 étudiants, les résultats sont les suivantes :

8 15 4 10 7 7 10 13 17 15  
3 6 13 12 9 10 16 11 10 8

#### Partie (I)

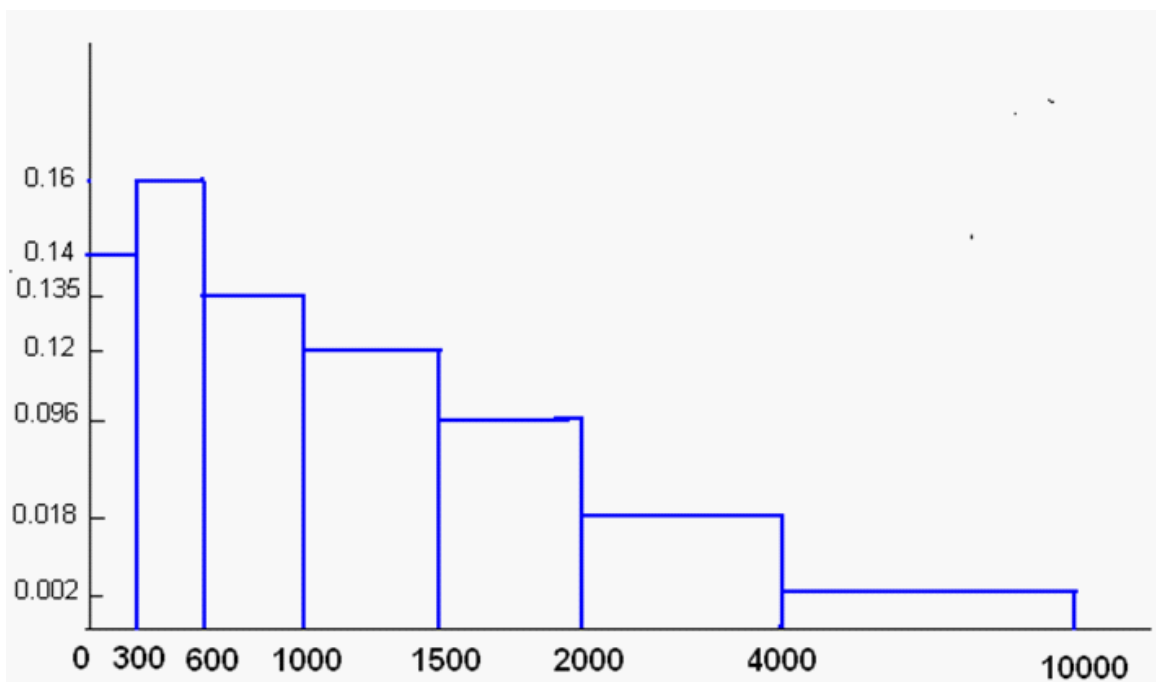
- Quelle est la nature de la variable ?
- Tracer le diagramme correspond à cette variable ainsi que le polygone des effectifs.
- Calculer la moyenne arithmétique  $\bar{X}$  et la médiane  $Me$  de cette distribution.

#### Partie (II)

- Ranger ces résultats en 5 classes de même amplitude de 4 points. (On prendre la borne inférieure de la première classe = 0).
- Tracer l'histogramme de cette distribution et la courbe des fréquences cumulées.
- Calculer les trois quartiles de cette distribution.
- En posant :  $y_i = (x_i - 10)/4$ , calculer  $\bar{Y}$ ,  $V(Y)$  et  $\sigma_Y$ . En déduire  $\bar{X}$ ,  $V(X)$  et  $\sigma_X$ .
- Expliquer pourquoi la moyenne et la médiane obtenues en (II) sont différentes de celle obtenue en (I).

#### Ex 2

L'histogramme ci-dessous décrit la distribution de la revenue annuelle ( en DA) chez les employés d'une entreprise en 1970.



- la fréquence relative de chacune des sept classes.
- Calculer la moyenne, les trois quartiles du revenue annuelle des employés.
- Interpréter les résultats du calculs.

#### Ex 3

Soit le tableau de contingence suivant :

$\begin{matrix} \backslash & Y \\ X & \backslash \end{matrix}$	0	1	2
1	$k^2$	$k$	$\frac{k}{2}$
2	$k^2 + 1$	$k^2$	$\frac{3k}{2}$

- 1) Déterminer la distribution marginale des effectifs de  $X$  et la distribution marginale des effectifs de  $Y$  en fonction de  $k$  ( $k > 0$ ).
- 2) Sachant que la moyenne conditionnelle de  $Y$ , pour  $x = 1$  vaut  $\frac{4}{7}$  ( $\bar{Y}_{x=1} = \frac{4}{7}$ ), déterminer la valeur de  $k$  dans ce cas.
- 3) Après avoir calculer la valeur de  $k$ 
  - Quelle est l'effectif total de la distribution du couple  $(X, Y)$ .
  - Déterminer la variance marginale de  $X$  et la variance marginale de  $Y$ .

## 4.7 Examen de Rattrapage 2011

### Rattrapage Statistique 2011 (1er LMD MI)

#### Exercice 1

Une série statistique  $X$ , dont les caractères représentent le nombre d'heures de sommeil durant la nuit d'un étudiant pendant une période de 50 jours, est donnée par le tableau suivant :

Caractère	1	2	3	4	5	6	7	8	9	10	11	12
Effectifs $n_i$	1	0	2	3	4	13	12	8	5	1	0	1

- 1) Quelle est le type du caractère étudié ?
- 2) Déterminer le mode, la médiane et la moyenne arithmétique de  $X$ .
- 3) On répartit ces valeurs en classes déterminées par les intervalles :  
 $[0.5; 3.5]$      $[3.5; 6.5]$      $[6.5; 9.5]$      $[9.5; 12.5]$ 
  - a) Calculer l'effectif, la fréquence, la fréquence cumulée croissante et décroissante de chaque classe.
  - b) Tracer l'histogramme et le polygone des effectifs.
  - c) Déterminer la classe modale et la classe médiane.
  - d) Calculer la médiane, la moyenne et la variance obtenues en faisant ce regroupement.

#### Sol

1)  $X$  : le nombre d'heures de sommeil de type **quantitatif discret**.

2) \* Le mode :  $\max n_i = 13 \implies Mo = 6$ .

\* La médiane :  $n = 50 = 2 \times 25 \implies p = 25 \implies Me = \frac{x_{25} + x_{26}}{2} \implies Me = \frac{7+7}{2} = 7$ .

\* La moyenne arithmétique :  $\bar{X} = \frac{1}{n} \sum_{i=1}^k n_i \times x_i = \frac{1+5+12+20+78+84+64+45+10+12}{50} = \frac{331}{50} = 6.62$

3) a)

$X$	$[0.5; 3.5]$	$[3.5; 6.5]$	$[6.5; 9.5]$	$[9.5; 12.5]$	<b>Total</b>
$n_i$	3	20	25	2	50
$f_i$	0.06	0.4	0.5	0.04	1
$c_i$	2	5	8	11	/
$F(x) \uparrow$	0.06	0.46	0.95	1	/
$F(x) \downarrow$	1	0.94	0.54	0.04	/

b) L'histogramme et le polygone des effectifs

:/swp55/temp/graphics/swp0000<sub>A1.pdf</sub>

Histogramme

c) \* La classe modale c'est la classe la plus élevée :  $[6.5; 9.5]$

\* la classe médiane c'est la classe qui vérifie :  $F(Me) = 0.5$

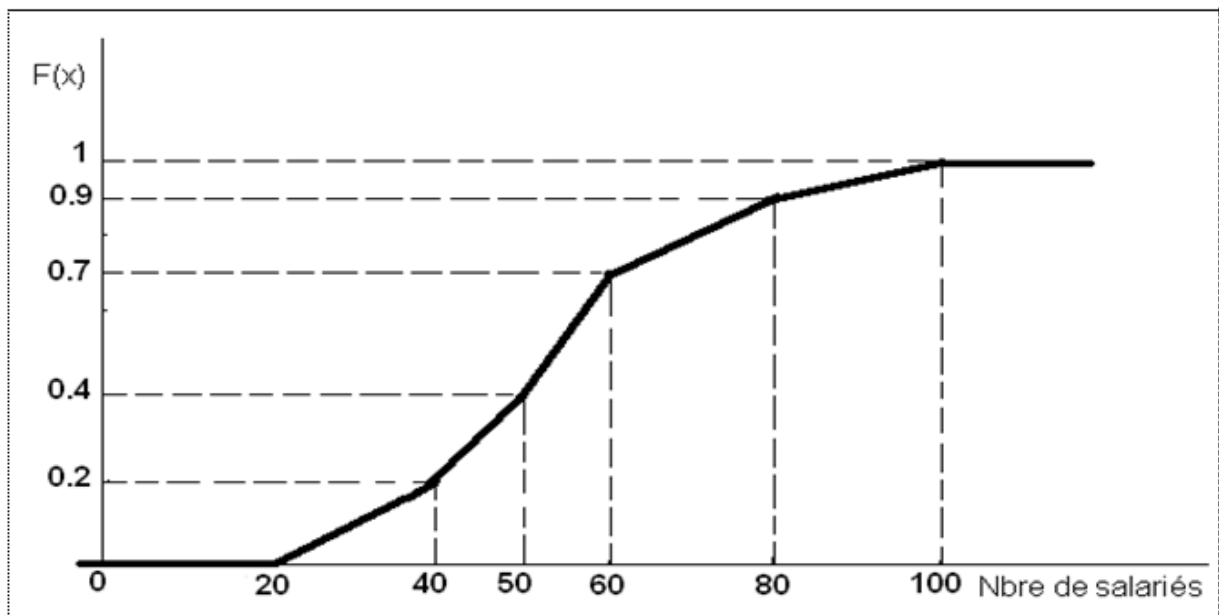
$0.46 = F(6.5) \leq 0.5 = F(Me) < 0.95 = F(9.5) \implies Me \in [6.5; 9.5]$

d) \* la médiane :  $Me \in [6.5; 9.5]$ , utilisons la méthode d'interpolation linéaire

$Me = 6.5 + 3 \times \frac{0.5-0.46}{0.95-0.46} = 6.74$

\* la moyenne :  $\bar{X} = \sum_{i=1}^k f_i \times c_i = 0.06 \times 2 + 0.4 \times 5 + 0.5 \times 8 + 0.04 \times 11 = 6.56$

\* la variance :  $V(X) = \sum_{i=1}^k f_i \times c_i^2 - (\bar{X})^2 = [0.06 \times (2)^2 + 0.4 \times (5)^2 + 0.5 \times (8)^2 + 0.04 \times (11)^2] - (6.56)^2 = 47.08 - (6.56)^2 = 4.04$ .



### Exercice 2

Sur un échantillon des entreprises industrielles, on a réalisé une enquête sur le nombre de salariés de ces entreprises.

On a réparti ces entreprises selon le nombre de salariés en 5 classes (limites de classes : 20, 40, 50, 60, 80 et 100). On a alors pu obtenir le graphique suivant des fréquences relatives cumulées (Fonction de répartition  $F(x)$ ).

- 1) Calculer la fréquence relative de chacune des cinq classes.
- 2) Calculer la moyenne et la médiane de nombre de salariés de l'échantillon.
- 3) Dessiner l'histogramme des fréquences.

**Sol**

1) Dressons d'abord le tableau statistique :

Classes	Fréquences cumulées $F(x)$	Amplitudes $a_i$	Centres $c_i$	Fréquences $f_i$	Fréquences rectifiés $f_i' = 10 \times \frac{f_i}{a_i}$
[20; 40[	0.2	20	30	0.2	0.1
[40; 50[	0.4	10	45	0.2	0.2
[50; 60[	0.7	10	55	0.3	0.3
[60; 80[	0.9	20	70	0.2	0.1
[80; 100[	1	20	90	0.1	0.05
<b>Total</b>	/	/	/	1	/

2) \* la moyenne :

$$\bar{X} = \sum_{i=1}^k f_i c_i = 0.2 \times 30 + 0.2 \times 45 + 0.3 \times 55 + 0.2 \times 70 + 0.1 \times 90 \implies \bar{X} = 54.5.$$

\* la médiane : On l'obtient en prenant l'abscisse du point d'ordonnée 0, 50 sur la courbe des fréquences cumulées précédente.

- La valeur de la médiane s'obtient, dans la classe médiane [50; 60[, par interpolation linéaire entre les points (50; 0, 4) et (60; 0, 7) :

$$Me = 50 + (60 - 50) \frac{0.50 - 0.4}{0.7 - 0.4} = 53.33$$



3) L'histogramme des fréquences : Construction de l'histogramme : Comme les classes ici ne sont pas de même amplitude, alors, les hauteurs des rectangles de l'histogramme doivent être proportionnelle aux fréquences rectifiées

$$f_i^{\lambda} = f_i \times \frac{\min(a_i)}{a_i} = 10 \times \frac{f_i}{a_i}$$

:/swp55/temp/graphics/swp0000\_43.pdf

Histogramme

### Exercice 3

Soit le tableau de contingence suivant :

	Y			
X	\	[0, 5[	[5, 12[	[12, 20[
[0, 8[		20	10	0
[8, 16[		10	20	10
[16, 20[		0	10	20

Où  $X$  représente les notes de statistiques et  $Y$  les notes d'analyse.

- 1) Déterminer les deux distributions marginales.
- 2) Calculer la moyenne et la variance des notes de statistiques.
- 3) Calculer la moyenne des notes d'analyse pour les étudiants qui ont obtenus au moins 8/20 en statistiques.

**Sol**

	Y				
X	\	[0, 5[	[5, 12[	[12, 20[	$X(n_{i.})$
[0, 8[		20	10	0	30
[8, 16[		10	20	10	40
[16, 20[		0	10	20	30
$Y(n_{.j})$		30	40	30	$n = 100$

1) \*Distribution marginale de  $X$  :

X	[0, 8[	[8, 16[	[16, 20[	<b>Total</b>
$n_{i.}$	30	40	30	100
$c_i$	4	12	18	/

\* Distribution marginale de  $Y$  :

Y	[0, 5[	[5, 12[	[12, 20[	<b>Total</b>
$n_{.j}$	30	40	30	100
$c_j$	2.5	8.5	16	/

2) la moyenne et la variance des notes de statistiques

$$\bar{X} = \frac{1}{n} \sum_{i=1}^3 n_{i.} \times c_i = \frac{4 \times 30 + 12 \times 40 + 16 \times 30}{100} = 10.8$$

$$V(X) = \frac{1}{n} \sum_{i=1}^3 n_{i.} \times c_i^2 - (\bar{X})^2 = \frac{(4)^2 \times 30 + (12)^2 \times 40 + (16)^2 \times 30}{100} - (10.8)^2$$

$$= 139.2 - 116.64 = 22.56$$

3)

$Y_{x \in [8,16[}$	$[0, 5[$	$[5, 12[$	$[12, 20[$	<b>Total</b>
$n_{2j}$	10	20	10	$n_{2.} = 40$
$c_j$	2.5	8.5	16	/

$$\bar{Y}_{x \in [8,16[} = \frac{1}{n_{2.}} \sum_{j=1}^3 n_{2j} \times c_j = \frac{10 \times 2.5 + 20 \times 8.5 + 10 \times 16}{40} = 8.875$$

$Y_{x \in [16,20[}$	$[0, 5[$	$[5, 12[$	$[12, 20[$	<b>Total</b>
$n_{3j}$	0	10	20	$n_{3.} = 30$
$c_j$	2.5	8.5	16	/

$$\bar{Y}_{x \in [16,20[} = \frac{1}{n_{3.}} \sum_{j=1}^3 n_{3j} \times c_j = \frac{0 \times 2.5 + 10 \times 8.5 + 20 \times 16}{30} = 13.5$$

## 4.8 Examen 2010

### Examen de Statistique 2010

#### Ex 1

On relevé les revenus annuels d'un échantillon de 20 personnes. Les valeurs sont les suivantes (en centaines de DA) :

180 195 230 215 180 185 220 230 190 190  
180 220 250 185 180 190 230 220 180 190

#### I)

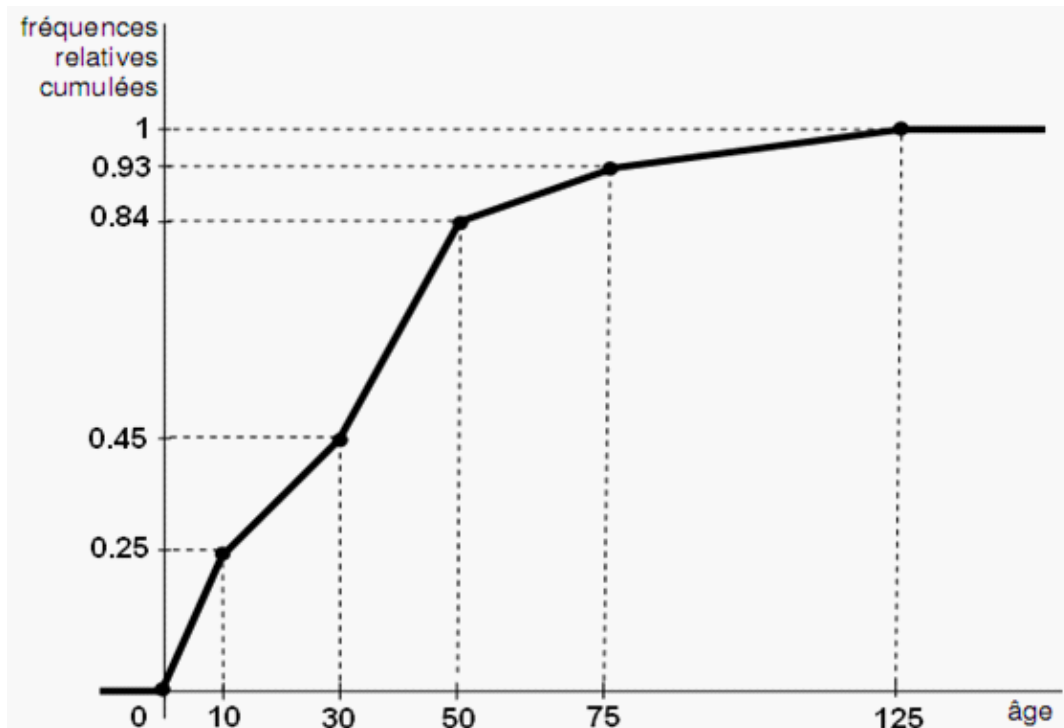
- Quelle est la nature de la variable ?
- Tracer le diagramme correspondant à cette variable ainsi que le polygone des effectifs.
- Calculer la moyenne arithmétique  $\bar{X}$  et la médiane  $Me$  de cette distribution.

#### II)

- Ranger ces résultats en classe de même amplitude de  $l$ .
- Tracer l'Histogramme de cette distribution et la courbe des fréquences cumulées.
- Calculer les 1<sup>er</sup>, 2<sup>ème</sup> et 3<sup>ème</sup> quartile de cette distribution.
- En posant :  $y_i = (x_i - 180)/15$ , calculer  $\bar{Y}$ ,  $V(Y)$  et  $\sigma_Y$ . En déduire  $\bar{X}$ ,  $V(X)$  et  $\sigma_X$ .
- Expliquer pourquoi la moyenne et la médiane obtenues en II) sont différentes de celle obtenue en I).

#### Ex 2

Sur un échantillon d'habitants d'une certaine région, on a observé l'âge de ces habitants. On a réparti ces individus en 5 classes d'âge (limites de classe : 0, 10, 30, 50, 75 et 125). On a alors pu obtenir le graphique suivant des fréquences relatives cumulées (fonction de répartition) :



- Calculer la fréquence relative de chacune des cinq classes.
- Calculer la moyenne et la médiane de l'âge des individus de l'échantillon.

3) Dessiner l'histogramme des fréquences.

**Ex 3**

$X$  et  $Y$  représentent deux variables statistiques qui prennent les valeurs 0 et 1.

Pour  $n$  individus, on recueille  $a$  observations du type (0,0),  $b$  observations du type (0,1),  $c$  observations du type (1,0),  $d$  observations du type (1,1).

Cette distribution statistique est donnée par le tableau suivant :

$X \setminus Y$	0	1
0	$a$	$b$
1	$c$	$d$

Donner, en fonction de  $a, b, c, d, n$  ( $n = a + b + c + d$ ) :

- 1) Les distributions marginales des fréquences.
- 2) Les moyennes et variances marginales.
- 3) Les moyennes et variances conditionnées.
- 4) La covariance  $Cov(X, Y)$ .
- 5) Le coefficient de corrélation linéaire  $r(X, Y)$ .

## 4.9 Examen de Rattrapage 2010

### Rattrapage de Statistique 2010

#### Ex 1

On a relevé les nombres d'étudiants de 1<sup>er</sup> année dans 20 groupes de TD, les résultats sont les suivants :

36    38    28    34    36    44    26    34    32    36  
 30    36    30    36    34    36    38    40    32    38

#### I)

- Quelle est la nature de la variable ?
- Tracer le diagramme correspondant à cette variable ainsi que le polygone des effectifs.
- Calculer la moyenne arithmétique  $\bar{X}$  et la médiane  $Me$  de cette distribution.

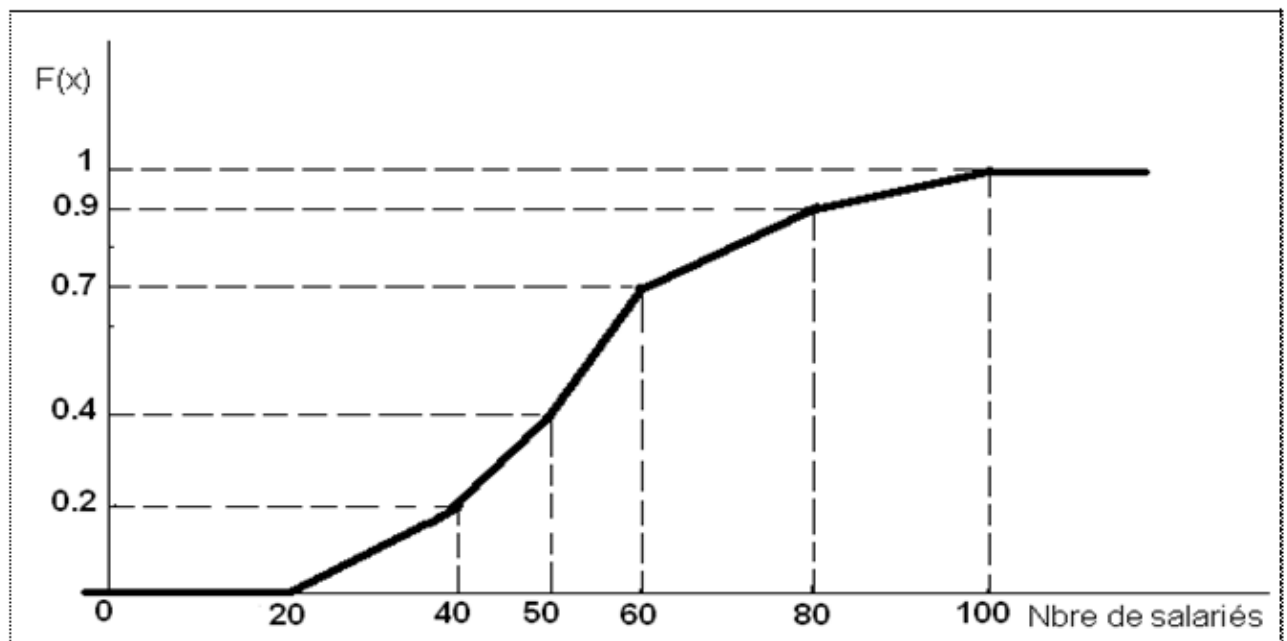
#### II)

- Ranger ces résultats en classe d'amplitude de 4 étudiants.
- Tracer l'Histogramme de cette distribution et la courbe des fréquences cumulées.
- Calculer les 1<sup>er</sup> et 2<sup>ème</sup> quartile de cette distribution.
- En posant :  $y_i = \frac{x_i - 26}{4}$ , calculer  $\bar{Y}$ ,  $V(Y)$  et  $\sigma_Y$ . En déduire  $\bar{X}$ ,  $V(X)$  et  $\sigma_X$ .
- Expliquer pourquoi la moyenne et la médiane obtenues en II) sont différentes de celle obtenue en I).

#### Ex 2

Sur un échantillon des entreprises industrielles, on a réalisé une enquête sur le nombre de salariés de ces entreprises.

On a réparti ces entreprises selon le nombre de salariés en 5 classes (limites de classes : 20, 40, 50, 60, 80 et 100). On a alors pu obtenir le graphique suivant des fréquences relatives cumulées (Fonction de répartition  $F(x)$ ).



- Calculer la fréquence relative de chacune des cinq classes.

- 2) Calculer la moyenne et la médiane de nombre de salariés de l'échantillon.
- 3) Dessiner l'histogramme des fréquences.

**Ex 3**

Soit le tableau de contingence suivant :

$X \backslash Y$	$[0, 5[$	$[5, 12[$	$[12, 20[$
$[0, 8[$	20	10	0
$[8, 16[$	10	20	10
$[16, 20[$	0	10	20

Où  $X$  représente les notes de statistiques et  $Y$  les notes d'analyse.

- 1) Déterminer les deux distributions marginales.
- 2) Calculer la moyenne et la variance des notes de statistiques.
- 3) Calculer la moyenne des notes d'analyse pour les étudiants qui ont obtenus au moins 8/20 en statistiques.
- 4) Calculer le coefficient de corrélation  $r(X, Y)$ .

## 4.10 correction examen de Rattrapage 2010

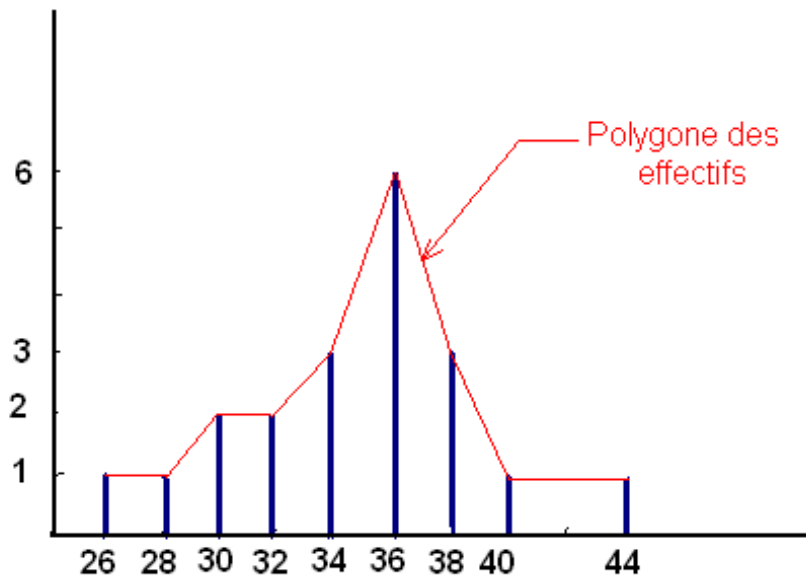
### Exercice 1

I)

a) le caractère étudié :  $X$  le nombre d'étudiants, son type : quantitatif discret.

Nombre d'étudiants $x_i$	26	28	30	32	34	36	38	40	44	Total
Effectifs $n_i$	1	1	2	2	3	6	3	1	1	20

- Le diagramme correspondant à cette variable est le diagramme en bâton :



c) \* La moyenne :

$$\begin{aligned}\bar{X} &= \frac{1}{n} \sum_{i=1}^n n_i x_i = \frac{1}{20} (26 \times 1 + 28 \times 1 + 30 \times 2 + 32 \times 2 + 34 \times 3 + 36 \times 6 + 38 \times 3 + 40 \times 1 + 44 \times 1) \\ &= \frac{694}{20} = 34.7\end{aligned}$$

\* La médiane : - On doit d'abord ordonner la série statistique dans l'ordre croissant :

Comme  $n = 20$  est pair alors  $n = 2p = 2 \times 10$

$$\Rightarrow p = 10 \Rightarrow Me = \frac{x_{10} + x_{11}}{2} = \frac{36 + 36}{2} = 36.$$

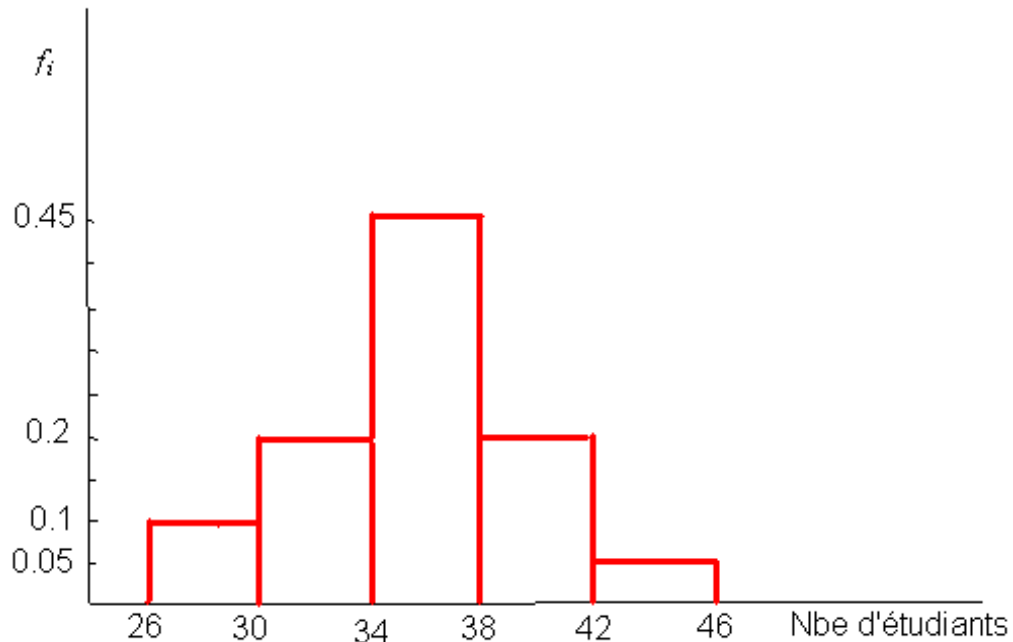
II)

a) Rangement des résultats en classe de même amplitude  $a_i = 4$  :

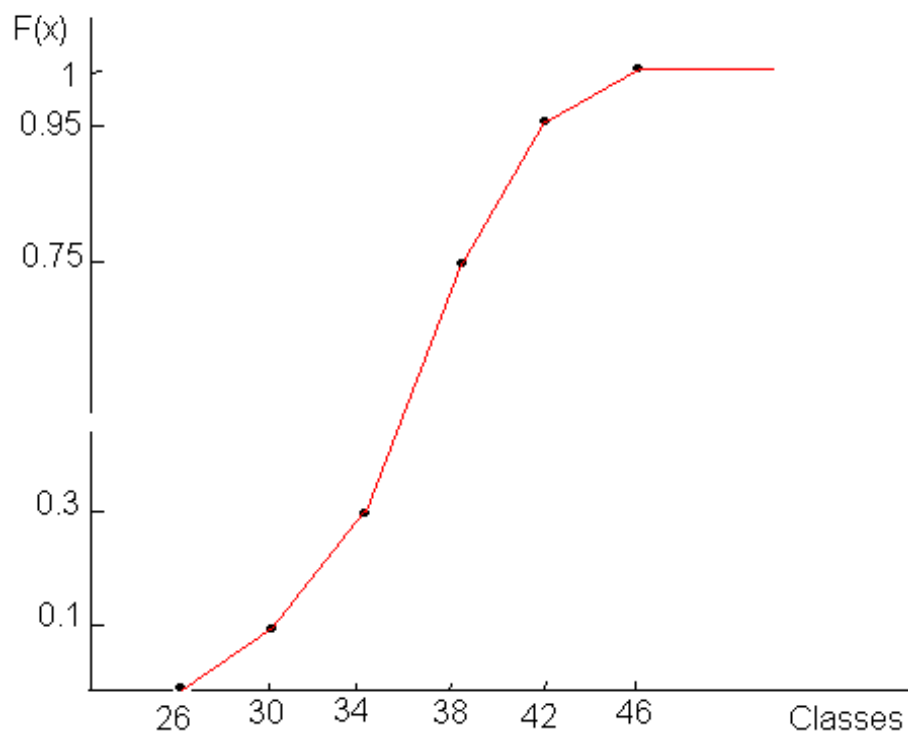
- Dressons d'abord le tableau statistique en calculons les centres de classes  $c_i$ , les fréquences  $f_i$ , les amplitudes  $a_i$  et les fréquences cumulées (ou fonction de répartition  $F(x)$ ) :

nombre d'étudiants	[26; 30[	[30; 34[	[34; 38[	[38; 42[	[42; 46[	Total
Effectifs $n_i$	2	4	9	4	1	20
Centres $c_i$	28	32	36	40	44	/
Fréquences $f_i$	0.1	0.2	0.45	0.2	0.05	1
F. cum $F(x)$	0.1	0.3	0.75	0.95	1	/

b) \* L'Histogramme : comme les classes ici sont de même amplitude, alors les hauteurs des rectangles de l'histogramme doivent être proportionnelles à la fréquence  $f_i$  :



\* La courbe des fréquences cumulées : on joint les points ayant pour abscisses la limite supérieure des classes et pour ordonnées les fréquences relatives cumulées correspondant à la classe considérée.





c) Les quartiles : pour calculer les quartiles, on a besoin des fréquences cumulées  $F(x)$ .

le premier quartile correspond à la valeur 0.25 de la fonction de répartition i.e

$$F(Q_1) = 0.25$$

$$Q_1 \in [30; 34[ \implies Q_1 = 30 + (34 - 30) \frac{0.25-0.1}{0.3-0.1} = 33.0$$

la médiane correspond à la valeur 0.50 de la fonction de répartition i.e  $F(Me) = 0.50$

$$Me \in [34; 38[ \implies Me = 34 + (38 - 34) \frac{0.5-0.3}{0.75-0.3} = 35.778$$

d)

nombre d'étudiants	[26; 30[	[30; 34[	[34; 38[	[38; 42[	[42; 46[	Total
Fréquences $f_i$	0.1	0.2	0.45	0.2	0.05	1
$y_i = \frac{c_i-26}{4}$	0.5	1.5	2.5	3.5	4.5	/

- Calculons  $\bar{Y}$ ;  $V(Y)$  et  $\sigma_Y$

$$\bar{Y} = \sum_{i=1}^k y_i f_i = 0.1 \times 0.5 + 0.2 \times 1.5 + 0.45 \times 2.5 + 0.2 \times 3.5 + 0.05 \times 4.5 = 2.4$$

$$V(Y) = \sum_{i=1}^k y_i^2 f_i - \bar{Y}^2 =$$

$$(0.1 \times (0.5)^2 + 0.2 \times (1.5)^2 + 0.45 \times (2.5)^2 + 0.2 \times (3.5)^2 + 0.05 \times (4.5)^2) - (2.4)^2 = 6.75 - 5.76 = 0.99 \approx 1$$

$$\text{D'où } \sigma_Y = \sqrt{V(Y)} \approx 1$$

- Déduisons  $\bar{X}$ ;  $V(X)$  et  $\sigma_X$

$$\text{Comme } y_i = \frac{x_i-26}{4} \implies \bar{Y} = \frac{\bar{X}-26}{4} \implies \bar{X} = 4\bar{Y} + 26$$

$$\text{D'où } \bar{X} = 4(2.4) + 26 = 35.6$$

$$V(X) = 16V(Y) = 16$$

$$\text{D'où } \sigma_X = \sqrt{V(X)} = 4$$

e) La différence tient au fait qu'en regroupant les valeurs en classes, on perd un peu d'information et aussi au fait qu'à l'intérieur des classes, les modalités ne sont pas équi-réparties.

### Exercice 2

1) Dressons d'abord le tableau statistique :

Classes	Fréquences cumulées $F(x)$	Amplitudes $a_i$	Centres $c_i$	Fréquences $f_i$	Fréquences rectifiés $f_i' = 10 \times \frac{f_i}{a_i}$
[20; 40[	0.2	20	30	0.2	0.1
[40; 50[	0.4	10	45	0.2	0.2
[50; 60[	0.7	10	55	0.3	0.3
[60; 80[	0.9	20	70	0.2	0.1
[80; 100[	1	20	90	0.1	0.05
<b>Total</b>	/	/	/	1	/

2) \* la moyenne :

$$\bar{X} = \sum_{i=1}^k f_i c_i = 0.2 \times 30 + 0.2 \times 45 + 0.3 \times 55 + 0.2 \times 70 + 0.1 \times 90 \implies \bar{X} = 54.5.$$

\* la médiane : On l'obtient en prenant l'abscisse du point d'ordonnée 0,50 sur la courbe des fréquences cumulées précédente.

- La valeur de la médiane s'obtient, dans la classe médiane [50; 60[, par interpolation linéaire entre les points (50; 0,4) et (60; 0,7) :

$$Me = 50 + (60 - 50) \frac{0.50 - 0.4}{0.7 - 0.4} = 53.33$$

3) L'histogramme des fréquences : Construction de l'histogramme : Comme les classes ici ne sont pas de même amplitude, alors, les hauteurs des rectangles de l'histogramme doivent être proportionnelles aux fréquences rectifiées

$$f_i^* = f_i \times \frac{\min(a_i)}{a_i} = 10 \times \frac{f_i}{a_i}$$

# Exercices Divers

---

**EX 1.** Soient les données suivantes relevées sur un échantillon de 200 étudiants de L1 de l'Université de Paris 1 Panthéon Sorbonne :

Taille (cm)	$n_i$	$F(x)$
[150; 160[		0,2
[160; 170[	60	
[170; 180[		0,6
[180; 190[	30	
[190; 200[		

1/ Compléter ce tableau.

2/ Calculer moyenne, variance et écart type de cette distribution.

3/ Déterminer la médiane.

4/ Représenter graphiquement la courbe cumulative de cette distribution.

**Corrigé**

**EX 2.** On donne la série statistique suivante :

$x_i$	0	1	2	3	4	5
$n_i$	13	32	24	$a$	$b$	7

1/ Déterminer les inconnues **a** et **b** sachant que la moyenne de la distribution vaut 1,97 et que l'écart type est de 1,42.

2/ Représenter le diagramme en bâtons de cette distribution.

**EX 3.** Les salaires de l'entreprise STEFRIC SA. se répartissent ainsi (en milliers d'euros) :

[1; 2[	[2; 4[	[4; 7[	[7; 8[	[8; 10[	[10; 12[
50	35	42	21	14	3

1/ De quelle nature est la variable «salaires de l'entreprise STERIC S.A.» ?

2/ Etablir la distribution de la variable.

3/ Déterminer les amplitudes des différentes classes de salaires.

4/ Calculer le mode, la moyenne et la médiane de cette distribution.

5/ Calculer l'intervalle interquartile de cette distribution.

6/ Quel salaire « minimum » perçoivent les 10% des salariés les mieux payés de l'entreprise STERIC SA. ?

7/ Représenter graphiquement cette distribution.

EX 4. La série suivante représente le nombre de pièces non-conformes par jour, dans une entreprise :122-111-154-98-93-67-134-167-123-142-132-151-127-119-137-130-127-135-187-165-161-

151-143-148-132-127-99-132-139-100-136-132-127-167-118-116-83-77.

- 1) Classer ces données dans des intervalles de même amplitude.
- 2) Tracer le diagramme différentiel et le diagramme intégral.
- 3) Calculer la médiane et déterminer la classe modale.
- 4) Calculer les moyennes arithmétique, géométrique, harmonique, et quadratique.
- 5) Etudier l'asymétrie et la forme de cette série.
- 6) Calculer le coefficient de variation.

**Corrigé**

1) Rangeons les données par ordre croissant

67-77-83-93-98-99-100-111-116-118-119-122-123-127-127-127-127-130-132-132-132-134-135-136-137-139-142-143-148-151-151-154-161-165-167-167-187

Utilisons la règle empirique de Sturges pour déterminer le nombre de classes :

$$k = 1 + 0.3 \ln(n) = 1 + 3.3 \log(38) \cong 13$$

Il en résulte un nombre de classe égal à 13

D'autre part l'étendue

$$E = \max(x_i) - \min(x_i) = 187 - 67 = 120.$$

Par la suite l'amplitude de chaque classe sera :

$$a = \frac{E}{k} \cong 10.$$

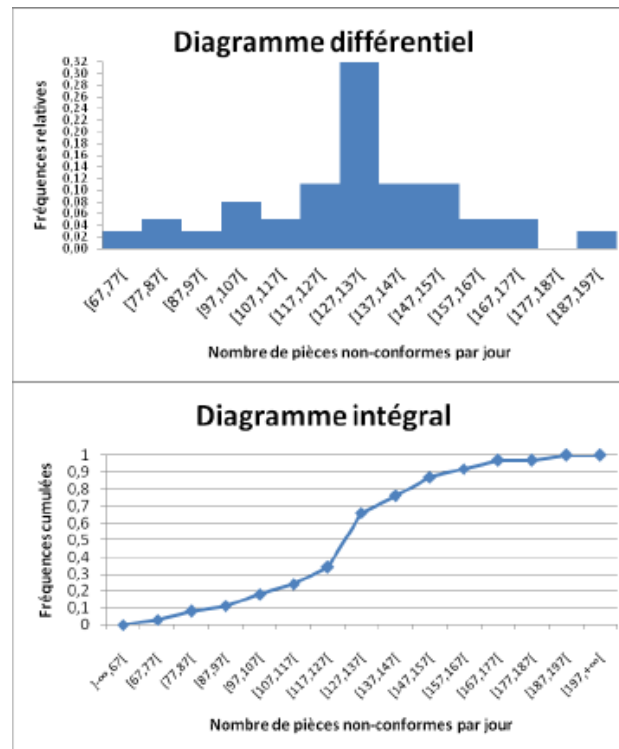
D'où la présentation des données sous forme d'un tableau statistique :

$[e_i; e_{i+1}[$	$n_i$	$e_i$	$f_i$	$f_i^*$	$f_i \cdot e_i$	$f_i \cdot e_i^2$	$\frac{f_i}{e_i}$	$f_i \cdot \ln(e_i)$	$f_i \cdot (e_i - \bar{X})^3$	$f_i \cdot (e_i - \bar{X})^4$
[67,77[	1	72	0,03	0,03	1,89	136,42	$3,65 \cdot 10^{-4}$	0,11	-5317,25	312015,97
[77,87[	2	82	0,05	0,08	4,32	353,89	$6,42 \cdot 10^{-4}$	0,23	-6071,53	295562,05
[87,97[	1	92	0,03	0,11	2,42	222,74	$2,86 \cdot 10^{-4}$	0,12	-1522,92	58906,37
[97,107[	3	102	0,08	0,18	8,05	821,37	$7,74 \cdot 10^{-4}$	0,37	-1862,41	53413,89
[107,117[	2	112	0,05	0,24	5,89	660,21	$4,70 \cdot 10^{-4}$	0,25	-343,07	6408,46
[117,127[	4	122	0,11	0,34	12,84	1566,74	$8,63 \cdot 10^{-4}$	0,51	-68,84	597,52
[127,137[	12	132	0,32	0,66	41,68	5502,32	$2,39 \cdot 10^{-3}$	1,54	0,73	0,96
[137,147[	4	142	0,11	0,76	14,95	2122,53	$7,41 \cdot 10^{-4}$	0,52	152,69	1728,47
[147,157[	4	152	0,11	0,87	16,00	2432,00	$6,93 \cdot 10^{-4}$	0,53	1020,09	21748,29
[157,167[	2	162	0,05	0,92	8,53	1381,26	$3,25 \cdot 10^{-4}$	0,27	1617,01	50644,63
[167,177[	2	172	0,05	0,97	9,05	1557,05	$3,06 \cdot 10^{-4}$	0,27	3713,02	153422,00
[177,187[	0	182	0,00	0,97	0,00	0,00	0,00	0,00	0,00	0,00
[187,197[	1	192	0,03	1,00	5,05	970,11	$1,37 \cdot 10^{-4}$	0,14	6067,68	372070,29
$\Sigma$	38		1,00		130,68	17726,63	$7,99 \cdot 10^{-3}$	4,85	-2614,79	1326518,92

2) Diagramme différentiel / Diagramme intégral

3)

■ Médiane :



$$Me = e_i + a_i \left( \frac{0.5 - F_i \uparrow}{F_{i+1} \uparrow - F_i \uparrow} \right) = e_7 + a_7 \left( \frac{0.5 - F_6 \uparrow}{F_7 \uparrow - F_6 \uparrow} \right) = 127 + 10 \left( \frac{0.5 - 0.34}{0.66 - 0.34} \right) = 132$$

■ Classe modale : [127, 137[

4)

■ Moyenne Arithmétique :  $\bar{X} = \sum_{i=1}^{13} f_i c_i = 130.68$

■ Moyenne Géométrique :  $G = \sqrt[n]{\prod_{i=1}^{13} c_i^{n_i}}$

Ou encore :  $\ln(G) = \sum_{i=1}^{13} f_i \ln(c_i) = 4.85 \Leftrightarrow G = e^{4.85} = 127.74$

■ Moyenne Harmonique :

$$H = \frac{n}{\sum_{i=1}^{13} \frac{n_i}{c_i}} \Leftrightarrow H^{-1} = \sum_{i=1}^{13} \frac{f_i}{c_i} = 7.99 \times 10^{-3} \Leftrightarrow H = 125.165$$

■ Moyenne Quadratique :

$$Q = \sqrt{\sum_{i=1}^{13} f_i c_i^2} = \sqrt{17726.63} = 133.14$$

La relation suivante sera toujours vérifiée :  $H \leq G \leq \bar{X} \leq Q$

**5) Etude de l'asymétrie :**

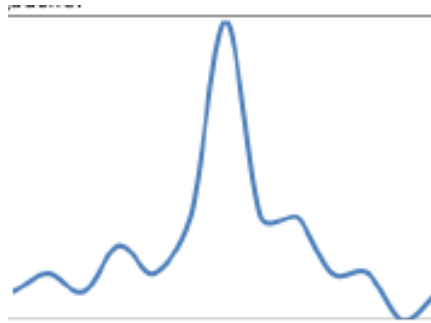
Calculons le coefficient de dissymétrie  $\gamma_1 = \frac{\mu_3}{\sigma^3}$

Avec  $\mu_3 = \sum_{i=1}^{13} f_i \cdot (c_i - \bar{X})^3 = -2614,79$  le 3<sup>ème</sup> moment centré d'ordre 3.

Rappelons que l'écart-type  $\sigma = \sqrt{V(X)}$  Où

$$V(X) = \sum_{i=1}^{13} f_i \cdot (c_i - \bar{X})^2 = \sum_{i=1}^{13} f_i c_i^2 - (\bar{X})^2 = 17726.63 - (130.68)^2 = 649.37$$

Par la suite  $\sigma = 25.48$ .



Ce qui donne :  $\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{-2614.79}{(25.48)^3} = -0.16 < 0$ .

En effet distribution étalée vers la droite ou dissymétrique à droite donc une queue de distribution étalée vers la gauche.

#### Etude de l'aplatissement :

Calculons le coefficient de Fisher  $\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$  avec

$$\mu_4 = \sum_{i=1}^{13} f_i \cdot (c_i - \bar{X})^4 = 1326518.92$$

$$\text{Par la suite } \gamma_2 = \frac{1326518.92}{(649.37)^2} - 3 = 0.15 > 0$$

Par conséquent la distribution est plus affilée que la distribution normale. Elle est dite leptocurtique.

#### 6) Calculons le coefficient de variation

$$C_v = \frac{\sigma}{\bar{X}} = 0.19 = 19\%.$$

**EX 5.** La saisie des données statistiques a été effectuée. On a mesuré le tour de poitrine des 30 étudiants et les résultats exprimés en cm sont les suivantes :

85-79-96-95-95-80-82-73-84-80-93-92-74-85-92  
92-80-84-78-80-93-97-86-99-84-83-90-76-83-82

- 1) Ordonner ces données.
- 2) Etablir un tableau statistique des effectifs, des fréquences et des fréquences cumulées. Tracer la courbe représentative de la fonction de répartition.
- 3) Déterminer la médiane.

#### Corrigé

1)

73-74-76-78-79-80-80-80-80-82-82-83-83-84-84  
84-85-85-86-90-92-92-92-93-93-95-95-96-97-99

2)

■ Variable statistique quantitative discrète : Taille

$x_i$	$n_i$	$f_i$	$F_i \uparrow$
1	5	0.167	0.167
2	14	0.467	0.633
3	8	0.267	0.900
4	3	0.100	1
Total	30	1	

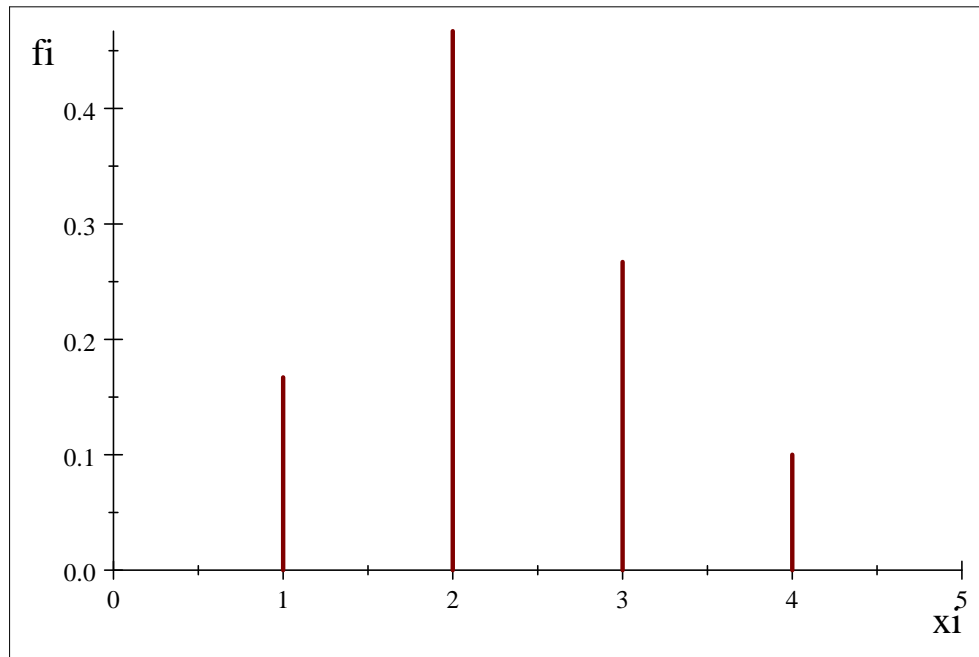
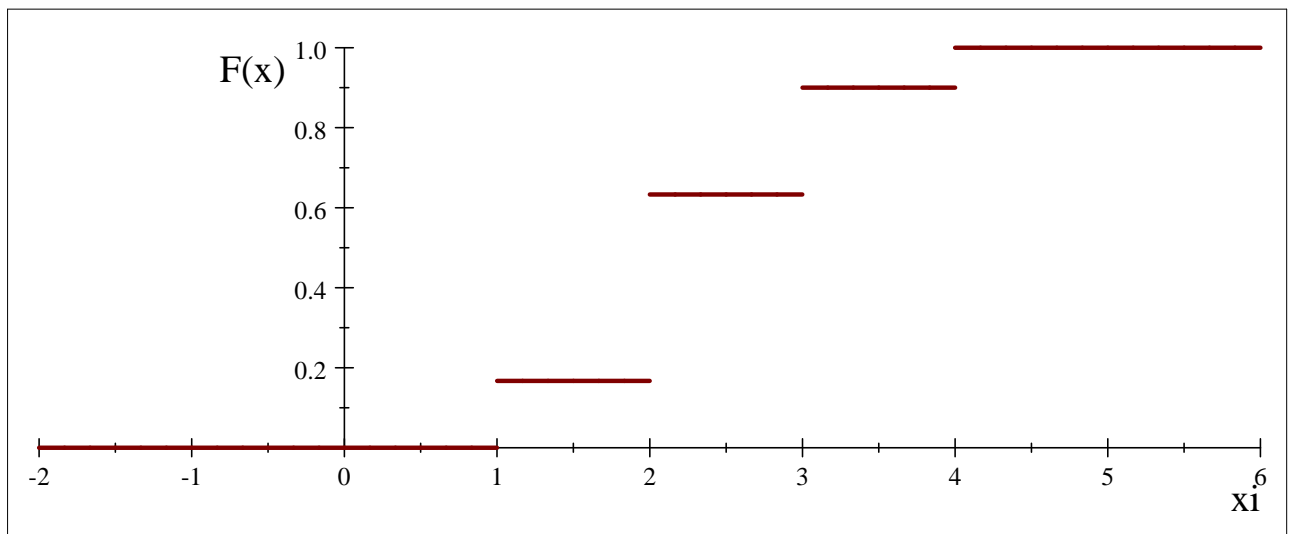
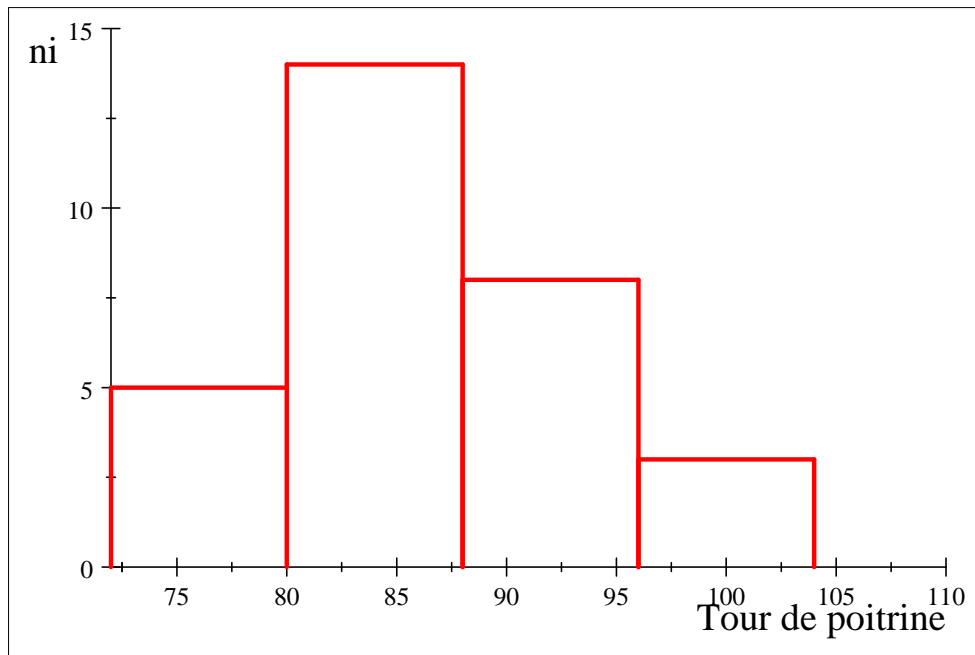


Diagramme en bâtons

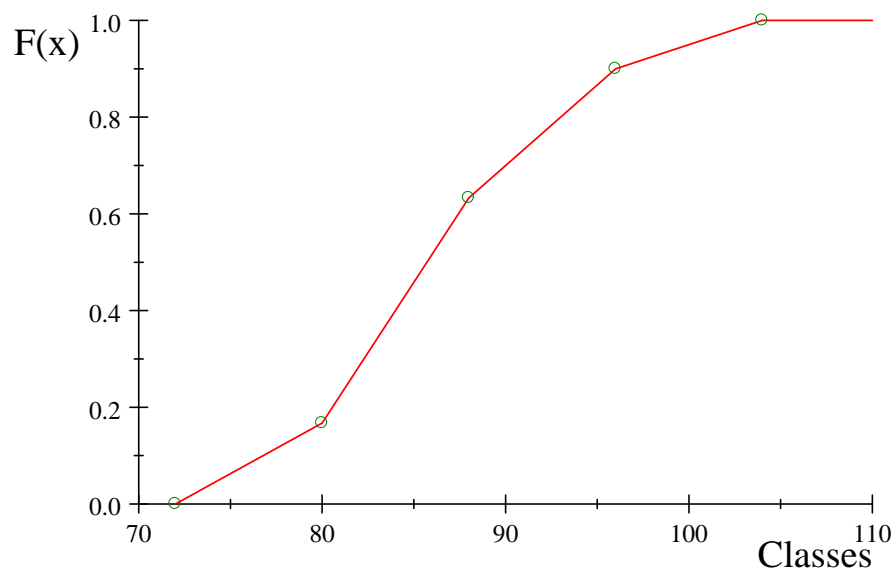


■ Variable statistique quantitative continue : Tour de poitrine

$[e_{i+1}, e_i[$	$n_i$	$f_i$	$F_i \uparrow$
$[72, 80[$	5	0.167	0.167
$[80, 88[$	14	0.467	0.633
$[88, 96[$	8	0.267	0.900
$[96, 104[$	3	0.100	1
<b>Total</b>	30	1	



Histogramme



## 3) La Médiane

$$Me = e_i + a_i \left( \frac{0.5 - F_i \uparrow}{F_{i+1} \uparrow - F_i \uparrow} \right) = e_2 + a_2 \left( \frac{0.5 - F_1 \uparrow}{F_2 \uparrow - F_1 \uparrow} \right) = 80 + 8 \left( \frac{0.5 - 0.167}{0.633 - 0.167} \right) = 85.717$$



# Bibliographie

---

- [1] PILLER. A., Statistique Descriptive, éditions Premium, 2004
- [2] GRAIS. B., Statistique descriptive, coll. « Éco sup », 3<sup>e</sup> éd., Dunod, 2004.
- [3] GRAIS. B., Méthodes statistiques, coll. « Éco sup », 4<sup>e</sup>éd., Dunod, 2006.
- [4] DODGE. Y., Statistique, Dictionnaire encyclopédique, Springer Verlag, 2004.
- [5] PY. B., Exercices corrigés de statistique descriptive, 2<sup>e</sup> édition, Economica, 1994.
- [6] SAPORTA. G., Probabilités, Statistique et Analyse des Données. Technip, 1990.
- [7] DELMAS. J.F., Introduction au calcul des probabilités et a la statistique. Polycopié ENSTA, 2005.
- [8] BOULEAU. N., Probabilités de l'ingénieur. Variables aléatoires et simulation. Hermann, 1986.
- [9] MORINEAU. A., CHATELINY. M. ET AL., L'analyse statistique des données : apprendre, comprendre et réaliser avec Excel, Ellipses, 2005.
- [10] FELLER. W., An introduction to probability theory and its applications, volume 1. Wiley and Sons, Inc., 3rd edition, 1968.