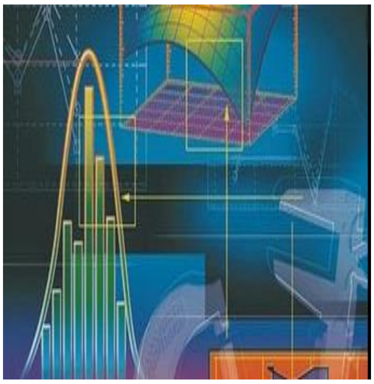
République Algérienne Démocratique et populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Akli Mohand Oulhadj - Bouira
Faculté des sciences et sciences appliquées
Département de Génie des Procédés

Polycopié de cours

Matière : Probabilités & Statistiques

Niveau : 2éme année licence génie des procédés



Présenté par :

Dr. BELLACHE Dihia

Année universitaire : 2022/2023

PREAMBULE

Fiche contacte

Intitulé du domaine : Sciences et Technologie ;

Année : $2^{ème}$ année ;

Intitulé de la matière : Maths 4 Probabilités et Statistiques ;

Annuel ou semestriel : Semestriel ;

Unité d'enseignement : UEM3 Méthodologie ;

Volume horaire global 45 heures : 1h30 /semaine Cours
1h30 /semaine TD ;

Chargé de la matière : BELLACHE Dihia ;

Nombre de crédits : 4.

Pré-requis

Présentation du module

Objectifs du module

Ce module permet aux étudiants de voir les notions essentielles da la probabilité et de la statistique, à savoir : les séries statistiques à une et à deux variables, la probabilité sur un univers fini et les variables aléatoires.

Partie I : STATISITQUES

CHAPITRE I. DEFINITIONS DE BASE

CHAPITRE II. SERIES STATISTIQUES A UNE VARIABLE

CHAPITRE III. SERIES STATISTIQUES A DEUX VARIABLES

Partie II Probabilités

CHAPITRE I. ANALYSE COMBINATOIRE

CHAPITRE II. CALCUL DES PROBABILITES

CHAPITRE III. CONDITIONNEMENT ET INDEPENDANCE

CHAPITRE IV. VARIABLES ALEATOIRES

CHAPITRE V. LOIS DE PROBABILITE DISCRETES USUELLES

Partie A : Statistiques

CHAPITRE I. DEFINITIONS DE BASE

I.1. QUELQUES DEFINITIONS

I.1.1. La statistique

a. Définition

Le mot statistique désigne à la fois un ensemble de données, d'observations et l'activité qui consiste dans leur recueil, leur traitement et leur interprétation.

Étymologie : « De l'allemand Staatskunde, dérivé de l'italien statista (homme d'État, statiste), la statistique représente l'ensemble des connaissances que doit posséder un homme d'État. » (1785).

b. Histoire de la statistique

- Recensements en Chine au XXIIIe siècle av. J.-C. ou en Égypte au XVIIIe av. J.-C, système de recueil se poursuivant jusqu'au XVIIe.
- Rôle prévisionnel des statistiques au XVIIIe siècle avec la construction des premières tables de mortalité avec Antoine Deparcieux, l'Essai sur les probabilités de la durée de vie humaine (1746).
- Rôle démographique au XIXe siècle, le Baron de Reiffenberg présentait en 1842 à l'Académie ses calculs rétrospectifs de population chez des peuples gaulois, d'après des chiffres donnés par Jules César dans sa conquête des gaules.
- Premiers textes connus sur le calcul des hasards (ou des chances) au XVIe siècle avec Cardan et au XVIIe siècle avec Galilée.
- Début officiel avec Pascal, Fermat et Huyguens au XVIIe siècle.
- Tournant au XVIIIe siècle avec Montmort (combinatoire), Bernoulli (loi des grands nombres) puis De Moivre et Laplace (traitement analytique des probabilités et théorèmes limites).
- Théorie des ensembles et de la mesure par Borel et Lebesgue et calcul des probabilités par Lévy au XXe siècle
- Axiomatisation de la théorie des probabilités par Kolmogorov (1933).

c. Divers domaines d'application

- **Economie, assurance, finance** : Etudes quantitatives de marchés, prévisions économétriques, analyse de la consommation des ménages, taxation des primes d'assurances et de franchises, gestion de portefeuille, évaluation d'actifs financiers, ...etc.

- **Biologie, médecine** : Essais thérapeutiques, épidémiologie, dynamique des populations, analyse du génome, ...etc.
- **Sciences de la terre** : Prévisions météréologiques, exploration pétrolière, ...etc.
- **Sciences humaines** : Enquêtes d'opinion, sondages, étude de population, ...etc.
- **Sciences de l'ingénieur** : Contrôle qualité, sûreté de fonctionnement, évaluation des performances, ...etc.
- **Sciences de l'information** : Traitement des images et des signaux, reconnaissance de forme et parole, machine Learning.

d. But de la Statistique

Les données sont entachées d'incertitudes et présentent des variations pour plusieurs raisons :

- Le déroulement des phénomènes observés n'est pas prévisible à l'avance avec certitude ;
- Toute mesure est entachée d'erreur ;
- Seuls quelques individus sont observés...etc.

e. Classes des méthodes statistiques

- **Statistique descriptive** : Elle a pour but de **résumer l'information** contenue dans les données de façon synthétique et efficace par :
  - Représentations graphiques ;
  - Indicateurs de position, de dispersion et de relation ;
  - Régression linéaire.

Permet de dégager les caractéristiques essentielles du phénomène étudié et de suggérer des hypothèses pour une étude ultérieure plus poussée. Les probabilités n'ont ici qu'un rôle mineur.

- **Statistique inférentielle** : Elle a pour but de faire des **prévisions** et de **prendre des décisions** au vu des observations par :
  - Estimation paramétrique ;
  - Intervalles de confiance, tests d'hypothèse ;

Nécessite de définir des modèles probabilistes du phénomène aléatoire et savoir gérer les risques d'erreurs.

I.1.2. Population et individu

- **La population** est l'ensemble des éléments auxquels se rapportent les données étudiées.

Exemples I.1 :

Etudiants d'une université, production d'une usine, poissons d'une rivière, entreprise d'un secteur donnée.

- Dans une population donnée, chaque élément est appelé **individu** ou **unité statistique**.

Exemples I.2 :

Personne humaine, automobile, entreprise, pays,…etc.

I.1.3. Echantillon

Un échantillon est un sous-ensemble d'une population. Les données statistiques sont des mesures effectuées sur les individus de l'échantillon.

Exemple I.3 : Si l'échantillon est un groupe de TD des étudiants L2 en génie des procédés.

- Un individu est un étudiant ;
- La population peut être l'ensemble des étudiants L2 en génie de procédés, les étudiants de l'université de Bouira, les étudiants algériens…etc.
- Les variables étudiées peuvent être le sexe, la taille, la moyenne d'année, le nombre de cafés consommés,…etc.

I.1.4. Caractère (variable)

Pour étudier une population, le statisticien ne retient que les caractères qui l'intéressent.

Un caractère ou une variable statistique est une variable qui caractérise les individus de cette population.

Exemple I.4 :

Le poids, la taille, la couleur des yeux,… etc.

Les modalités sont les différentes situations susceptibles d'être prises par le caractère.

I.2. DIFFERENTS TYPES DE VARIABLES STATISTIQUES

Il existe deux grandes catégories de caractères : les caractères qualitatifs et les caractères quantitatifs.

I.2.1. Caractère qualitatif

Un Caractère est dit qualitatif si ses différentes réalisations (modalités) ne sont pas numériques.

Exemple I.5 : La catégorie socio-professionnelle des individus d'une population donnée (artisan, ouvrier,…etc.) est un caractère qualitatif.

I.2.2. Caractère quantitatif

Un caractère est dit quantitatif lorsqu'il est intrinsèquement numérique. Une variable quantitative peut être une variable statistique discrète ou continue.

a. **La variable statistique discrète**

La variable statistique X est dite discrète, lorsqu'elle ne peut prendre que des valeurs isolées, discrètes : $x_1, x_2, \ldots, x_n$ (où $x_1 < x_2 < \ldots < x_n$).

Exemple I.6 :

Le nombre d'enfants d'une famille, le nombre de buts marqués lors d'une rencontre de football, . . . etc., sont des variables (caractères) quantitatives discrètes.

b. La variable statistique continue,

La variable statistique X est dite continue, lorsqu'elle peut prendre n'importe quelle valeur d'un intervalle $[\alpha, \beta] \subset \mathbb{R}$.

Exemple I.7 :

La durée d'une conversation téléphonique, le revenu, la taille, le taux de natalité sont des variables continues.

Dans ce cas l'intervalle des valeurs possibles $[\alpha, \beta]$ est divisé en k intervalles : $[a_0, a_1[$, $[a_1, a_2[, \ldots, [a_{k-1}, a_k]$, (où $a_0 = \alpha < a_1 < a_2 < \ldots < a_k = \beta$), qui sont appelées des **classes**.

- Les valeurs $\mathbf{a_{k-1}, a_k}$ sont les **frontières** de la $k^{éme}$ classe.
- La valeur $\frac{a_{k-1}+a_k}{2}$ est le **centre de cette classe**.

Exemple récapitulatif :

On souhaite connaitre l'état de 100 maisons : Identifier la population, l'individu, le caractère et les modalités on utilisant les trois types de caractères.

a. Caractère qualitatif

- Population : Maisons (100) ;
- Individu : Une maison parmi ces 100 maisons ;

- Caractère : L'état de la maison ;
- Modalités : Petite, moyenne, grande ;

b. Caractère quantitatif discret

- Population : Maisons (100) ;
- Individu : Une maison parmi ces 100 maisons ;
- Caractère : Nombre de pièces ;
- Modalités : 1, 2, 3, 4, 5.

c. Caractère quantitatif continu

- Population : Maisons (100) ;
- Individu : Une maison parmi ces 100 maisons ;
- Caractère : Surface (notée S) ;
- Modalités : $S \in [60, 200]$ m².

CHAPITRE II. SERIES STATISTIQUES A UNE VARIABLE

En statistique, la liste de données est appelée série de données statistique. Etudier une série statistique correspond à l'étude d'un caractère (variable) dans une population.

Considérons une variable X observée sur une population de n individus. Si la variable X prend k modalités, le premier traitement des données brutes consiste à compter le nombre n_i d'individus qui présentent le i^{em} modalité, i = 1, 2, . . . , k.

II.1. EFFECTIF, FREQUENCE, POURCENTAGE

II.1.1. Effectif

L'effectif d'une modalité est le nombre d'individu de cette modalité. Généralement on note **n_i** l'effectif d'une modalité. C'est le nombre de fois que cette valeur apparait dans la liste. **L'effectif total** est la somme des effectifs de toutes les modalités. On le note souvent **N**, on a alors : $\mathrm{N} = \mathrm{n}_1 + \mathrm{n}_2 + \mathrm{n}_3 + \cdots \mathrm{n}_\mathrm{k}$, En utilisant la notation sigma : $\boldsymbol{N} = \sum_{i=1}^{k} \boldsymbol{n_i}$.

Exemple II.1 (Cas quantitatif discret) :

Dans une classe de 29 étudiants, le professeur a relevé les notes suivantes : 2−3−3− 4− 5 − 6−6− 7− 7−7−8−8−8−8−8−9−9−9−9−9−9−10−10−11−11−11−13−13−15. Afin d'y voir plus clair, il regroupe les notes dans une **série statistique A** représentée dans le tableau d'effectifs suivant.

Observations	**Notes (x_i)**	**Effectifs (n_i)**
1	2	1
2	3	2
3	4	1
4	5	1
5	6	2
6	7	3
7	8	5
8	9	6
9	10	2
10	11	3
11	13	2
12	15	1
		$\boldsymbol{N} = \sum_{i=1}^{12} \boldsymbol{n_i} = \mathbf{29}$

Exemple II.2 (Cas quantitatif continu) :

Le tableau suivent regroupe les classe d'une série statistique (**série B**), cette dernière représente les tailles de 120 sportifs exprimées en mètre divisée en classe d'amplitude 0.2m.

Classes (I_i)	Effectifs (n_i)
[1.60, 1.64 [	8
[1.64, 1.68 [	14
[1.68, 1.72 [	16
[1.72, 1.76 [	20
[1.76, 1.80 [	26
[1.80, 1.84 [	15
[1.84, 1.88 [	12
[1.88, 1.92 [	9
Total	$\mathbf{N} = \sum_{i=1}^{7} \boldsymbol{n_i} = 120$

II.1.2. Fréquence

On obtient la fréquence **(f_i)** pour une modalité en divisant l'effectif ou la proportion de la modalité **(n_i)** par l'effectif total **(N)**. On note $\boldsymbol{f_i = \frac{n_i}{N}}$; tel que : $\boldsymbol{\sum_{i=1}^{k} f_i = 1}$

Exemple II.3 (Cas quantitatif discret) :

Reprenons **l'exemple II.1** des notes obtenues de la classe de 29 étudiants :

Observation (i)	Notes Valeurs (x_i)	Etudiants Effectifs (n_i)	Fréquence (f_i)
1	2	1	$f_1 = \frac{n_1}{N} = \frac{1}{29} = 0.03448276$
2	3	2	$f_2 = \frac{n_2}{N} = \frac{2}{29} = 0.06896552$
3	4	1	$f_3 = \frac{n_3}{N} = \frac{1}{29} = 0.03448276$
4	5	1	$f_4 = \frac{n_4}{N} = \frac{1}{29} = 0.03448276$
5	6	2	$f_5 = \frac{n_5}{N} = \frac{2}{29} = 0.06896552$
6	7	3	$f_6 = \frac{n_6}{N} = \frac{3}{29} = 0.10344828$
7	8	5	$f_7 = \frac{n_7}{N} = \frac{5}{29} = 0.17241379$
8	9	6	$f_8 = \frac{n_8}{N} = \frac{6}{29} = 0.20689655$
9	10	2	$f_9 = \frac{n_9}{N} = \frac{2}{29} = 0.06896552$
10	11	3	$f_{10} = \frac{n_{10}}{N} = \frac{3}{29} = 0.10344828$
11	13	2	$f_{11} = \frac{n_{11}}{N} = \frac{2}{29} = 0.06896552$
12	15	1	$f_{12} = \frac{n_{12}}{N} = \frac{1}{29} = 0.03448276$
		$\boldsymbol{\sum_{i=1}^{12} n_i = 29}$	$\boldsymbol{\sum_{i=1}^{12} f_i = 1}$

II.1.3. Pourcentage

On obtient le pourcentage **(p_i)** pour une modalité en multipliant la fréquence **(f_i)** par 100, on note $\boldsymbol{p_i = \frac{n_i}{N} \times 100 = f_i \times 100}$; tel que : $\boldsymbol{\sum_{i=1}^{k} p_i = 100}$

Exemple II.4 (cas qualitatif) :

Selon The Nilson Report, Oct. 8, 1998, les 200 milliards achats par carte de crédits effectués aux USA pendant le premier semestre de l'année 1998 sont répartis selon la marque de la carte utilisée comme suit :

- 36 milliards achats avec la carte American express ;
- 2 milliards achats avec la carte Diners club ;
- 12 milliards achats avec la carte Discover ;
- 50 milliards achats avec la carte Master card ;
- 100 milliards achats avec la carte Visa.

Tableau des effectifs :

Modalités (Marque)	Effectifs (n_i)	Fréquences (f_i)	Pourcentage (P_i)
American express	36	0.18	18%
Diners club	2	0.01	1%
Discover	12	0.06	6%
Master card	50	0.25	25%
Visa	100	0.5	50%
Total	200	1	100%

II.2. EFFECTIF CUMULE, FREQUENCE CUMULEE.

II.2.1. Définition

Dans le cas d'une variable statistique discrète X, de modalités $x_1, x_2, x_3, \ldots, x_k$ et d'effectifs respectifs $n_1, n_2, n_3, \ldots n_k$, on appelle :

L'effectif cumulé ($N_i^{\nearrow}$) croissant associée à une valeur x_i de la variable X, le nombre d'individus n_i pour lesquels la variable X prend une valeur inférieure ou égale à x_i :

$\boldsymbol{N_i^{\nearrow} = \sum_{i=1}^{k} n_i = n_1 + n_2 + n_3 + \cdots + n_k}$; pour i=1, 2, 3,…, k.

L'effectif cumulé ($N_i^{\nearrow}$) dé**croissant** associée à une valeur x_i de la variable X, le nombre d'individus n_i pour lesquels la variable X prend une valeur supérieure ou égale à x_i :

$\boldsymbol{N_i^{\searrow} = \sum_{i=1}^{k} n_i = n_i + n_{i+1} + n_{i+2} + \cdots + n_k}$; pour i=1, 2, 3,…, k.

La fréquence cumulée croissante ($F_i^{\nearrow}$) de la valeur x_i est la fréquence de toutes les valeurs du caractère inférieures ou égales à x_i :

$\boldsymbol{F_i^{\nearrow} = \sum_{i=1}^{K} f_i = f_1 + f_2 + f_3 + \cdots + f_k = \frac{N_i^{\nearrow}}{N}}$; pour i=1, 2, 3,…, k.

La fréquence cumulée décroissante ($F_i^{\searrow}$) de la valeur xi est la fréquence de toutes les valeurs du caractère supérieures ou égales à x_i :

$$F_i^{\searrow} = \sum_{i=1}^{k} f_i = f_i + f_{i+1} + f_{i+2} + \cdots + f_k = \frac{N_i^{\searrow}}{N}\text{ ; pour i=1, 2, 3,…, k.}$$

Remarque II.1 :

Dans le cas d'une distribution continue, les données sont en général regroupées en classes. Les effectifs et les fréquences cumulés sont définis par rapport aux classes et non par rapport aux valeurs de la variable.

Exemples II.5 (Variable discrète) :

Reprenant la série statistique A de l'exemple II.1. Les effectifs cumulés croissants et les fréquences cumulées croissantes sont représentés dans le tableau suivant :

X_i	n_i	$N_i^{\nearrow}$	$N_i^{\searrow}$	f_i	$F_i^{\nearrow}$	$F_i^{\searrow}$
1	1	1	29	0.03	0.03448276	1
2	2	3	28	0.07	0.10344828	0.96551724
3	1	4	26	0.03	0.13793103	0.89655172
4	1	5	25	0.03	0.17241379	0.86206897
5	2	7	24	0.07	0.24137931	0.82758621
6	3	10	22	0.10	0.34482759	0.75862069
7	5	15	19	0.17	0.51724138	0.65517241
8	6	21	14	0.20	0.72413793	0.48275862
9	2	23	8	0.07	0.79310345	0.27586207
10	3	26	6	0.10	0.89655172	0.20689655
11	2	28	3	0.07	0.96551724	0.10344828
12	1	29	1	0.03	1	0.03448276
Total	29					

Exemple II.6 (Variable continue) :

Reprenant la sérié B de l'exemple II.2, les effectifs cumulés et les fréquences cumulées sont représentés dans ce tableau :

Classes (I_i)	n_i	$N_i^{\nearrow}$	$N_i^{\searrow}$	f_i	$F_i^{\nearrow}$	$F_i^{\searrow}$
[1.60, 1.64 [	8	8	120	0.07	0.06666667	1
[1.64, 1.68 [	14	22	112	0.12	0.18333333	0.93333333
[1.68, 1.72 [	16	38	98	0.13	0.31666667	0.81666667
[1.72, 1.76 [	20	58	82	0.17	0.48333333	0.68333333
[1.76, 1.80 [	26	84	62	0.22	0.7	0.51666667
[1.80, 1.84 [	15	99	36	0.12	0.825	0.3
[1.84, 1.88 [	12	111	21	0.1	0.925	0.175
[1.88, 1.92 [	9	120	9	0.07	1	0.075
Totale	120					

II.3. REPRESENTATIONS GRAPHIQUES

On peut associer à chaque aux tableaux statistiques des représentations graphiques qui ont pour objet essentiel de rendre visible, d'une façon globale, les données de ces tableaux. On distinguera d'une part, les graphiques basés sur les effectifs (ou les fréquences) appelés diagrammes différentiels qui mettent en évidence les différences d'effectifs (ou de fréquences) entre les modalités du caractère étudié et, d'autre part, les graphiques basés sur les effectifs cumulés (ou les fréquences cumulées) appelés diagrammes cumulatifs qui permettent de visualiser l'évolution des effectifs (ou des fréquences) cumulés.

II.3.1. Diagrammes différentiels

Un diagramme différentiel, appelé aussi graphique de distribution, peut prendre plusieurs formes, selon la nature des données.

II.3.1.1. Cas d'un caractère quantitatif et discret

Lorsque le caractère étudié est quantitatif et discret, on peut représenter la série statistique étudiée par :

a. Un diagramme en bâtons : La hauteur de chaque bâton est alors proportionnelle par rapport à l'effectif (ou à la fréquence) associé à chaque valeur.

Exemple II.7 :

Concéderons la série statistique de l'exemple II.1. La figure 1 suivante représente le diagramme en bâtons de la série A des notes des étudiants :

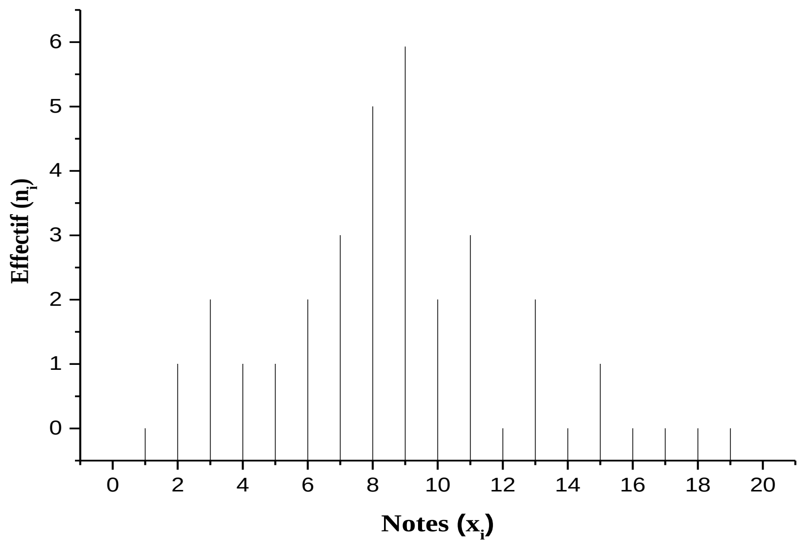


Figure II.1. Diagramme en bâtons des notes des étudiants (série A).

b. Un polygone des effectifs (ou des fréquences) : Il permet de représenter sous forme de courbe, la distribution des effectifs ou des fréquences. Le polygone des effectifs (ou des fréquences) est obtenu à partir du diagramme en bâtons des effectifs (ou des fréquences) en joignant par des segments de droite les sommets des bâtons.

Exemple II.8 :

On représente le polygone des effectifs de la série A (exemple II.1) dans la figure II.2 suivante :

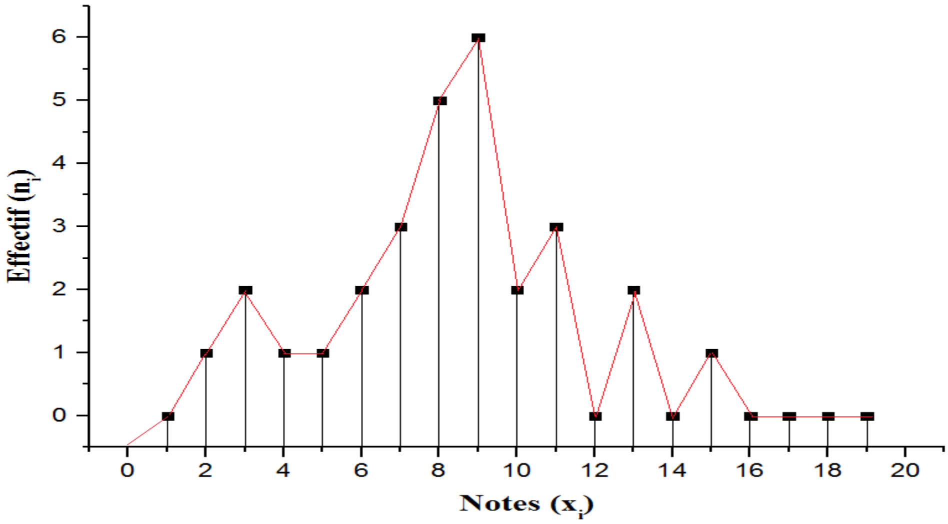


Figure II.2. Polygone des effectifs des notes des étudiants (série A).

II.3.1.2. Cas d'un caractère quantitatif et continu

Lorsque le caractère étudié est quantitatif et continu, et lorsque les modalités sont regroupées en classes, on peut représenter la série par :

a. Un histogramme : Un histogramme est un ensemble de rectangles juxtaposés dont les bases correspondent aux amplitudes des classes et dont les surfaces sont proportionnelles aux effectifs (ou aux fréquences) des classes. Nous admettons qu'il y a linéarité, c'est-à-dire que la distribution est uniforme à l'intérieur de chaque classe. On distingue deux cas :

1er cas : Les classes sont d'égale amplitude. On reporte en abscisse les bornes des classes et en ordonnée les effectifs (ou les fréquences). Dans ce cas, les hauteurs des rectangles sont proportionnelles aux effectifs (ou aux fréquences).

Exemple II.9 :

Reprenons la série B de l'exemple II.2, on obtient l'histogramme représenté dans la figure II.3 suivante :

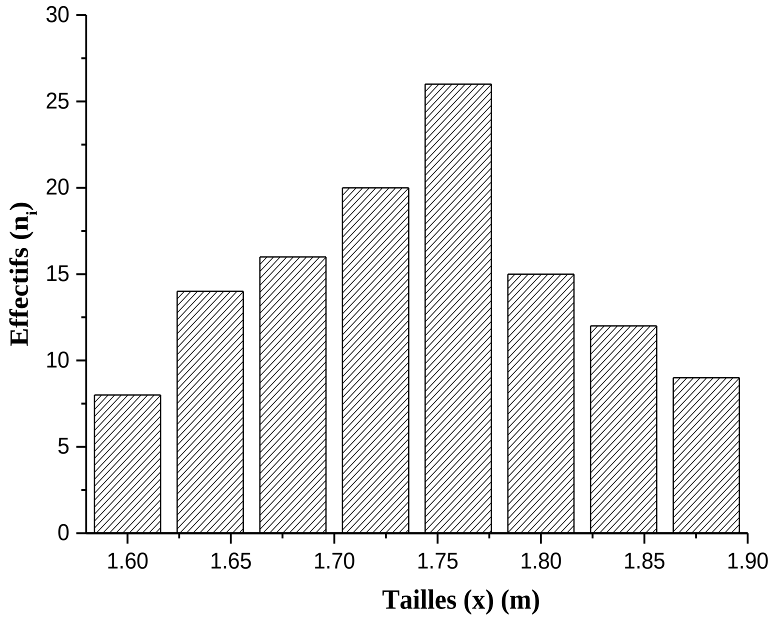


Figure II.3. Histogramme des effectifs des tailles des sportifs (Série B).

2ème cas : Les classes sont d'amplitudes inégales. Pour que la surface de chaque rectangle de l'histogramme reste proportionnelle à l'effectif (ou à la fréquence) correspondant, on reporte en abscisse les bornes des classes et en ordonnée **les densités d'effectifs** (ou de fréquences). La densité d'effectif (resp. de fréquence) d'une classe est définie comme le rapport de l'effectif n_i (resp. la fréquence f_i) à l'amplitude a_i de cette classe. Dans ce cas, la hauteur du rectangle n'est pas proportionnelle à l'effectif (ou la fréquence) de la classe mais à sa densité.

Exemple II.10 :

Le tableau suivant représente les salaires en euros des employés d'une entreprise :

Classes (I_i)	**Effectifs (n_i)**	**D_i (n_i/a_i)**	**Fréquences (f_i)**
[900, 1200 [	30	01	0.10714286
[1200, 1400 [	30	0.15	0.10714286
[1400, 1600 [	60	0.3	0.21428571
[1600, 1800 [	80	0.4	0.28571429
[1800, 2000 [	40	0.2	0.14285714
[2000, 2400 [	40	0.2	0.14285714
Total	**N=280**	**1.25**	**1**

Utilisant les données du tableau, on obtient l'histogramme représenté dans le tableau II.4 suivant :

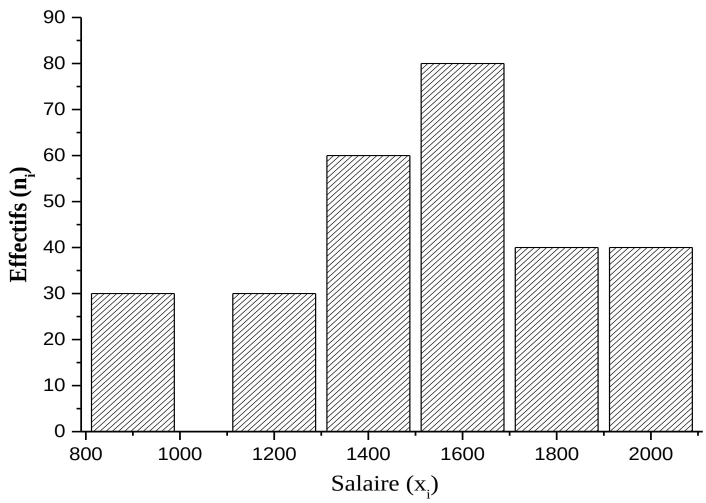


Figure II.4. Histogramme des salaires des employés d'une entreprise.

b. Un polygone des effectifs (et des fréquences) : Le polygone des effectifs (ou des fréquences) est obtenu en joignant, par des segments de droite, les milieux des côtés supérieurs de chaque rectangle de l'histogramme.

Exemple II.11 :

On représente le polygone des effectifs de la série B précédente (exemple II.2) dans la figure II.5 suivante :

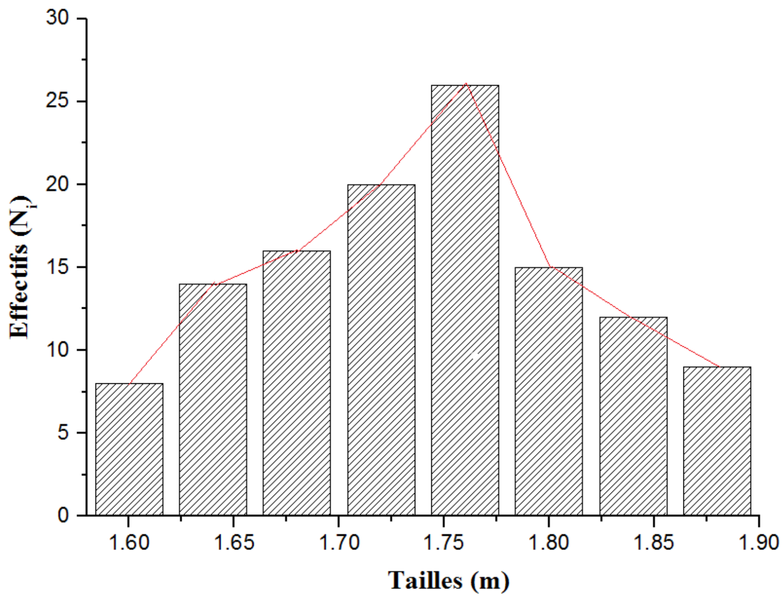


Figure II.5. Polygone des effectifs des tailles (série B).

II.3.1.3. Cas d'un caractère quantitatif

Enfin, lorsque le caractère est qualitatif, on peut représenter la série par :

- **Un diagramme circulaire (camemberts)** : La mesure de chaque secteur angulaire est proportionnelle à l'effectif associé.
- **Un diagramme en tuyaux d'orgue** : Chaque classe est représentée par un rectangle de même largeur et de longueur proportionnelle à l'effectif, donc à la fréquence.
- **Un diagramme en bandes** : Chaque classe est représentée par un rectangle de même largeur et de longueur proportionnelle à l'effectif, donc à la fréquence.

Exemple II.12 :

On représente le diagramme circulaire, le diagramme en tuyau d'orgue et le diagramme à bande des effectifs de la série C précédente (exemple II.4) dans la figure II.6 suivante :

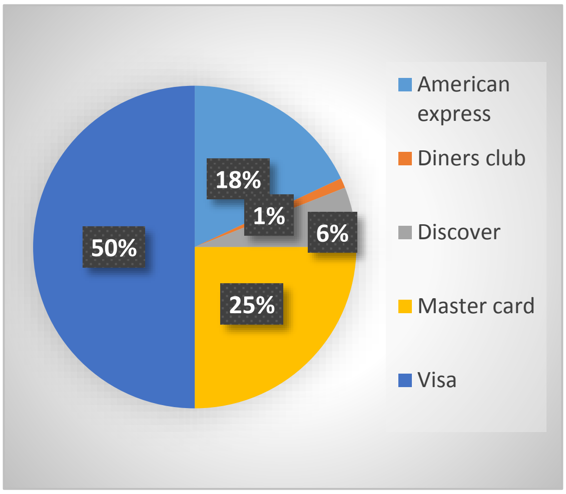


Diagramme circulaire (Fréquence)

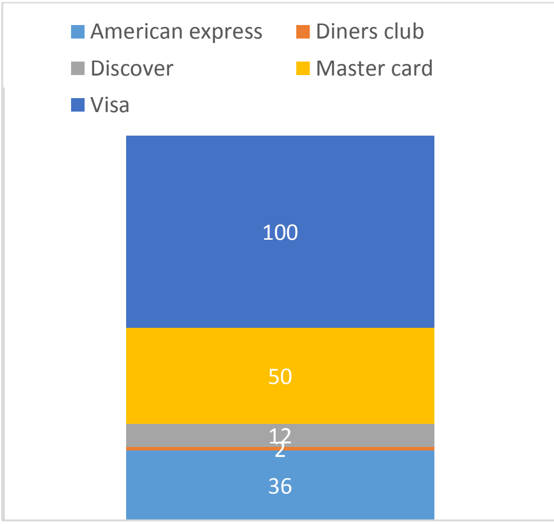


Diagramme à bande des effectifs

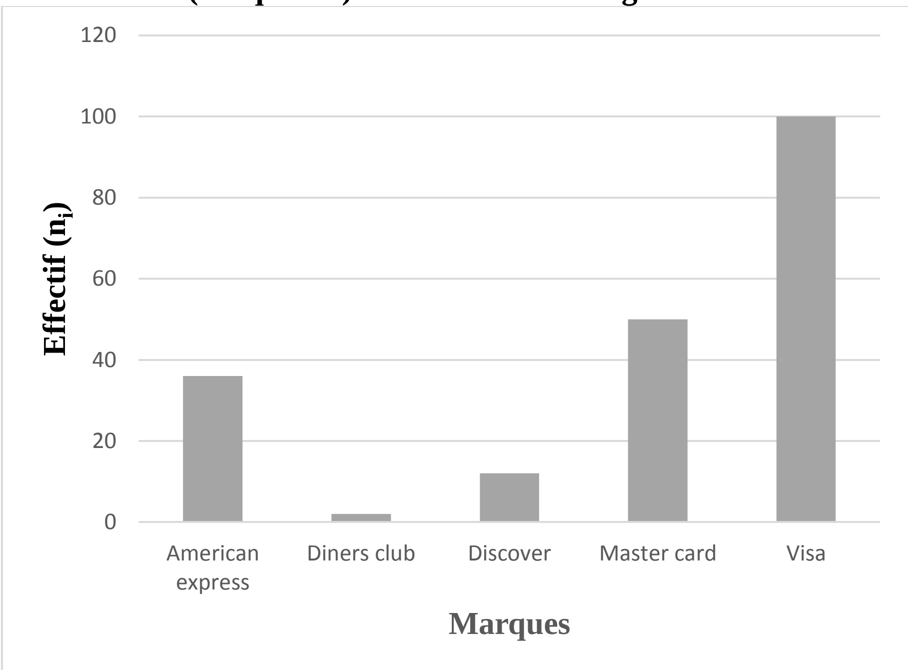


Diagramme en tuyau d'orgue des effectifs

Figure II.6. Représentation graphique d'un caractère qualitatif.

II.3.2. Courbes cumulatives

Le diagramme cumulatif (ou courbe cumulative) de la distribution d'une variable statistique est la représentation graphique des effectifs cumulés (ou des fréquences cumulées). Le diagramme cumulatif peut être croissant ou décroissant selon que l'on travaille avec les effectifs (ou les fréquences) cumulés croissants ou décroissants. Cette représentation graphique diffère suivant le type de la variable statistique.

II.3.2.1. Variable discrète

Le diagramme cumulatif d'une variable discrète se présente comme une courbe en escalier (i.e. constante par intervalle). Dans un premier temps, on place les points dont les abscisses sont les valeurs possibles de la variable, et dont les ordonnées sont égales aux effectifs cumulés (ou aux fréquences cumulées) correspondants. Ensuite, pour compléter le graphique, on trace des segments de droite horizontaux (les marches) pour chaque intervalle puisque, par définition, le cumul reste constant entre deux valeurs successives de la variable. Il important de noter que chaque segment de cette courbe en escalier est fermé à gauche et ouvert à droite (sauf le premier). Pour faciliter la lecture du graphique, on représente en plus des marches (trait plein), les verticales reliant ces marches, c'est-à-dire les contremarches (trait pointillé).

Exemple II.13 :

On représente dans la figure II.5 suivante les diagrammes cumulatifs des effectifs cumulés croissants et décroissants de la série en exemple II.1.

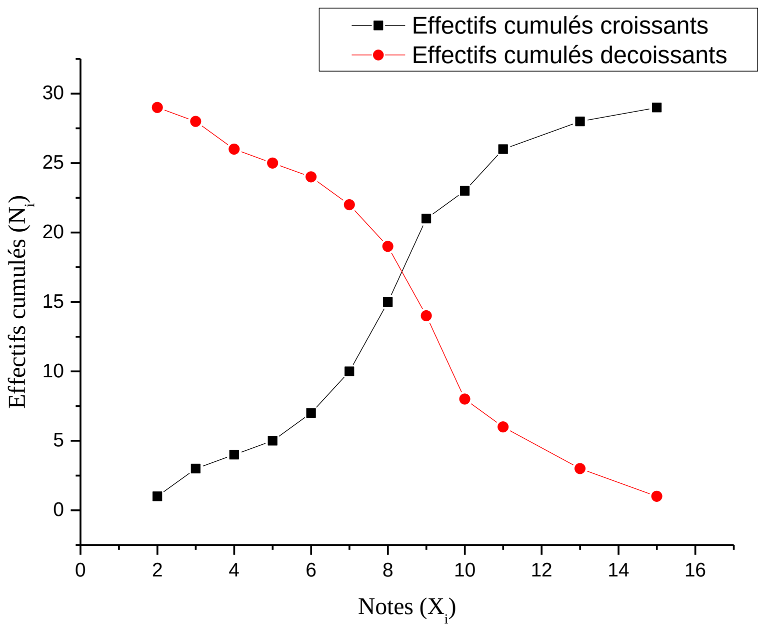


Figure II.5. Diagramme des effectifs cumulés des notes des étudiants.

II.3.2.3. Variable continue

Le diagramme cumulatif d'une variable continue prend la forme d'une courbe appelée polygone des effectifs cumulés (ou des fréquences cumulées).

- Le polygone des effectifs cumulés se construit en joignant, par des segments de droite, les points ayant pour abscisses les bornes supérieures des classes et pour ordonnées les effectifs cumulés croissants associés. On ajoute à ces points celui d'ordonnée nulle et d'abscisse égale à la borne inférieure de la première classe.

Exemple II.14 :

On représente dans les deux courbes des diagrammes effectifs cumules croissant et décroissant de la série en exemple II.2 dans la figure II.6.

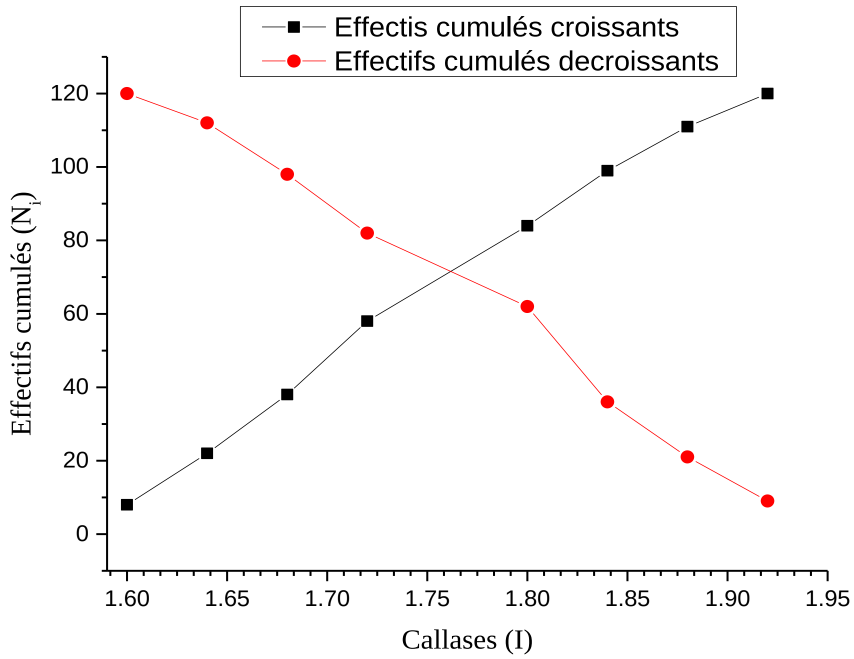


Figure II.6. Diagramme des effectifs cumulés des notes des étudiants.

Remarque II.2 :

Pour construire le polygone des fréquences cumulées, on remplace les effectifs par les fréquences dans ce qui précède.

La présence de classes d'amplitudes inégales n'entraine aucune modification concernant la construction des polygones des effectifs cumulés.

II.4. CARACTERISTIQUES DE POSITION

Ces caractéristiques nous informent sur l'ordre de grandeur des valeurs de la série statistique. On distingue d'une part, les caractéristiques de tendance centrale, comme la moyenne arithmétique, le mode et la médiane, qui permettent de déterminer une valeur centrale autour de laquelle les valeurs de la série ont tendance à se rassembler et, d'autre part, les caractéristiques de position non centrale, liée à un rang donné, comme les quantiles.

II.4.1. La moyenne arithmétique pondérée

a. Variable discrète

La moyenne arithmétique pondérée $\bar{x}$ d'une série statistique (x_i, n_i) i = 1, ..., k correspondant à une variable discrète est définie par :

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i$; pour des données isolées ;

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$; pour des données groupées par valeurs.

Exemple II.15 :

Dans la série de l'exemple II.1, la moyenne du contrôle est égale à :

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{238}{29} \simeq 8,207.$

b. Variable continue et données groupées par classes

Si la variable est continue et si les données sont groupées par classes, on prend pour valeur de x_i les centres de classes c_i. Ainsi, la moyenne arithmétique pondérée $\bar{x}$ s'écrit :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i$$

Exemple II.16 :

Dans la série de l'exemple II.10, la moyenne du contrôle est égale à :

$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i = \frac{460500}{280} = 1644.64$

II.4.2. Le mode

Le mode Mo d'une série est la valeur de la variable qui correspond au plus grand effectif (ou à la plus grande fréquence).

a. Variable discrète

Dans le cas d'une variable discrète, la détermination du mode est immédiate.

Exemple II.17 :

Reprenons le cas de l'exemple II.1, le mode de cette série est **Mo = 9**.

b. Variable continue et données groupées par classes

Si la variable est continue, et si les données sont groupées par classes, on parle plutôt de classe modale. On distingue deux cas :

1er cas : Les classes sont d'égale amplitude. La classe modale est la classe correspondant à l'effectif (ou à la fréquence) le plus élevé.

Exemple II.18 :

Reprenons la série statistique de l'exemple II.2, la classe modale est la classe [1.76, 1.80[.

2ème cas : Les classes dont l'amplitude inégale. Dans ce cas, la classe modale est la classe correspondant à la densité d'effectif (ou densité de fréquence) la plus élevée.

Exemple II.19 :

Reprenons la sérié statistique de l'exemple II.10, la classe modale est la classe [1600, 1800 [.

Remarque II.3 :

Une série statistique est dite uni-modale lorsqu'elle présente un seul mode (ou une seule classe modale), bimodale si elle en présente deux et plurimodale si elle en présente plus de deux. Il est possible qu'une série statistique ne présente aucun mode. Graphiquement, le mode correspond à l'abscisse d'un point d'ordonnée maximum du diagramme en bâtons ou du polygone des effectifs (ou des fréquences).

II.4.3. La médiane

La médiane Me d'une série statistique est la valeur telle que, la série étant ordonnée, il y ait autant d'observations rangées avant elle que d'observations rangées après elle. En d'autres termes, 50% des individus ont une valeur inférieure à la médiane et 50% une valeur supérieure. Sa détermination dépend du type de données. On distingue deux cas.

a. Variable discrète

Si l'effectif total n est impair, $n = 2p+1$ où $p \in \mathbb{N}^*$, la médiane Me est la valeur de rang p+1 lorsque l'on a ordonné les valeurs par ordre croissant. C'est une valeur observée de la série.

Si l'effectif total n est pair, $n = 2p$ où $p \in \mathbb{N}^*$, la médiane Me est comprise entre la valeur de rang p et la valeur de rang p+1. L'intervalle qui sépare ces deux valeurs est appelé intervalle médian et toute valeur de cet intervalle est une valeur médiane. On prend généralement pour Me, le centre de l'intervalle médian. Dans ces conditions, la médiane n'est pas une valeur observée de la série.

Exemple II.20 :

Reprenons la **série A** de l'exemple II.1 **:**

La médiane **Me= 8**.

b. Variable continue et données groupées par classes

La médiane Me est la valeur de la variable qui correspond à l'effectif cumulé n/2. On commence par chercher la classe médiane (l'intervalle où se trouve la médiane) à l'aide des effectifs cumulés. Si [x_{i-1}, x_i[désigne la classe médiane, l'effectif cumulé en x_{i-1} , noté N_{i-1} , est inférieur à n/2 et l'effectif cumulé en x_i , noté N_i , est supérieur à n/2. Pour déterminer Me on procède par interpolation linéaire en raison de l'hypothèse que la distribution est uniforme à l'intérieur de chaque classe. La figure II.7 suivante représente la méthode de calcul de la médiane par interpolation linéaire pour une variable continue.

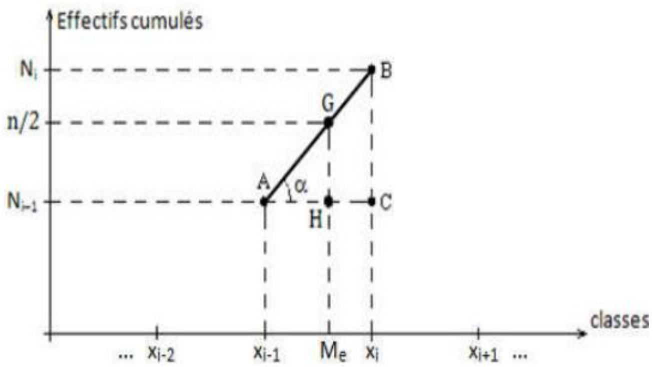


Figure II.7. Calcul de la médiane par interpolation linéaire pour une variable continue.

D'après la Figure II.7 ci-dessus :

$\tan\alpha = \frac{GH}{AH} = \frac{BC}{AC}$ D'où : $AH = AC\frac{GH}{BC}$, D'autre part : $AH = M_e - x_{i-1}$; $AC = x_i - x_{i-1}$; $GH = \frac{n}{2} - N_{i-1}$ et $BC = N_i - N_{i-1}$

Alors : $M_e - x_{i-1} = (x_i - x_{i-1})\frac{\frac{n}{2} - N_{i-1}}{N_i - N_{i-1}}$

Il en résulte :

$$\mathbf{M_e = (x_i - x_{i-1})\frac{\frac{n}{2} - N_{i-1}}{N_i - N_{i-1}} + x_{i-1}}$$

Remarque II.4 :

Graphiquement, la médiane Me peut être déterminée :

Soit à partir du polygone des effectifs cumulés comme l'abscisse du point d'ordonnée n/2,

Soit en repérant l'abscisse du point d'intersection du polygone des effectifs cumulés croissants et du polygone des effectifs cumulés décroissants.

Pour travailler avec le polygone des fréquences cumulées, on remplace les effectifs par les fréquences, et n/2 par 1/2 dans ce qui précède.

II.4.4. Les quartiles

Ils correspondent à des valeurs de la variable statistique qui partagent la série ordonnée en k parties égales :

- Si k = 4, les quantiles sont appelés quartiles. Il y a 3 quartiles, notés Q_1, Q_2, Q_3.
- Si k = 10, les quantiles sont appelés déciles. Il y a 9 déciles, notés D_1, D_2, ..., D_9.
- Si k = 100, les quantiles sont appelés centiles. Il y a 99 centiles, notés C_1, C_2, ..., C_{99}.

La détermination de ces caractéristiques est identique à celle de la médiane.

- Les quartiles Q_1, Q_2, Q_3 sont obtenus lorsque l'on a cumulé 25, 50, 75% de la population.
- Les déciles D_1, D_2, ..., D_9 sont obtenus lorsque l'on a cumulé 10, 20, ..., 90% de la population.
- Les centiles C_1, C_2, ..., C_{99} sont obtenus lorsque l'on a cumulé 1, 2, ..., 99% de la population.

Remarque II.5 :

Il va de soi qu'un grand nombre d'observations est nécessaire pour que les déciles et les centiles aient un sens.

II.5. CARACTERISTIQUES DE DISPERSION

Ces caractéristiques renseignent sur la dispersion des valeurs des données autour de la valeur centrale. Elles permettent aussi de comparer des séries entre elles. Les principales caractéristiques de dispersion sont : l'étendue, la variance, l'écart-type, le coefficient de variation.

II.5.1. Etendue

L'étendue, notée « **e** », d'une série statistique est la différence entre la plus grande et la plus petite des valeurs observées, dites valeurs extrêmes, de cette série :

$$\mathbf{e} = \mathbf{x}_{\mathbf{max}} - \mathbf{x}_{\mathbf{min}}$$

Cette mesure de la dispersion ne dépend que des valeurs extrêmes.

Exemple II.21 :

Reprenons la A de l'exemple II.1, on trouve l'étendue comme suit :

La note la plus élevée est 15, ($\mathbf{x_{max} = 15}$),

La plus basse est 2, ($\mathbf{x_{min} = 2}$)

L'étendue est donc égale à 13.

$e = x_{max} - x_{min} = 15 - 2 = 13$; $e = 13$.

II.5.2. L'étendue interquartile (écart interquartile)

De par la définition des quartiles, l'intervalle interquartile [Q_1, Q_3] contient 50 % des observations. Sa longueur, notée EIQ (Etendue Inter-Quartile), est un indicateur de dispersion :

$$\mathbf{EIQ = Q_3 - Q_1}$$

Notons que le calcul de l'étendue inter-quartile a l'avantage par rapport à celui de l'étendue d'écarter les valeurs extrêmes, souvent sans signification.

II.5.3. L'écart inter-décile

L'écart inter-décile (EID), est la différence entre le neuvième et le premier décile :

$$\mathbf{EID = D_9 - D_1}$$

Cet écart donne la longueur d'un intervalle contenant 80% des observations qui se trouvent dans la partie centrale de la population, quelles que soient les valeurs extrêmes observées.

II.5.4. Variance et écart-type

La variance d'une variable statistique x, notée V(x), est la moyenne arithmétique des carrées des écarts à la moyenne arithmétique $\bar{x}$:

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

Dans le cas d'une variable statistique continue, on ramène la valeur de chaque individu au milieu de sa classe d'affectation (x_i : centre de la $i^{éme}$ classe).

L'écart-type note σ d'une variable statistique (x_i) se définit comme la racine carrée de la variance :

$$\boldsymbol{\sigma} = \sqrt{\mathbf{V(x)}}$$

II.5.5. Coefficient de variation

Le coefficient de variation (CV) est le rapport de l'écart-type à la moyenne arithmétique :

$$\mathbf{CV = \frac{\sigma}{\bar{x}}}$$

C'est un nombre sans dimension. Souvent, il est exprimé en pourcentage. Plus le coefficient de variation est élevé, plus la dispersion autour de la moyenne arithmétique est élevée. Ce coefficient permet de comparer deux variables statistiques de natures différentes.

II.6. CARACTERISTIQUES DE FORME.

Ces caractéristiques indiquent la symétrie ou dissymétrie de la série des données, ainsi que son aplatissement. Les principales caractéristiques de forme sont les coefficients d'asymétrie et d'aplatissement de Fisher. Ces caractéristiques sont définies à partir de la notion de moment.

II.6.1. Les moments centrés

Le moment centré d'ordre r d'une distribution est égal à la moyenne arithmétique des puissances d'ordre r des écarts $(x_i - \bar{x})$.

- Pour des données isolées, le moment centré d'ordre r est :

$$\mu_r = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^r$$

- En revanche, pour des données groupées par valeurs, il s'écrit :

$$\mu_r = \frac{1}{n}\sum_{i=1}^{k} n_i(x_i - \bar{x})^r$$

- Enfin, pour des données groupées par classes, il devient :

$$\mu_r = \frac{1}{n}\sum_{i=1}^{k} n_i(c_i - \bar{x})^r$$

II.6.2. L'asymétrie

Le coefficient d'asymétrie de Fisher est le rapport, noté γ_1, entre le moment centré d'ordre 3 et le cube de l'écart-type, c'est-à-dire :

$$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

C'est un nombre sans dimension. Le coefficient d'asymétrie est nul si la répartition de l'échantillon ou de la distribution est symétrique. Il est positif si la répartition est étalée vers la droite, par contre il est négatif si la répartition est étalée vers la gauche.

II.6.3. L'aplatissement

Le coefficient d'aplatissement de Fisher est le rapport, noté γ_2, entre le moment centré d'ordre 4 et la puissance 4 de l'écart-type, rapport diminué de 3, c'est-à-dire :

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

C'est un nombre sans dimension. Le coefficient d'aplatissement est nul si la répartition des observations est méso-curtique, c'est-à-dire similaire à celle d'une loi normale. Le coefficient d'aplatissement est positif si la répartition est lepto-curtique, c'est-à-dire plus pointue que celle d'une loi normale de même moyenne et de même écart-type. À l'inverse, le coefficient d'aplatissement est négatif si la répartition est plati-curtique, c'est-à-dire plus aplatie que celle d'une loi normale de même moyenne et de même écart-type.

CHAPITRE III. SERIES STATISTIQUES A DEUX VARIABLES

Dans beaucoup de recherches statistiques, on ne s'intéresse pas qu'à un seul caractère mais à plusieurs en même temps. Quand on étudie deux caractères X et Y sur une population donnée, c'est en général parce qu'on cherche à savoir s'il existe un lien entre eux et qu'elle est l'intensité du lien.

Dans ce chapitre les méthodes décrites dans le cadre de la description des séries de données individuelles vont être étendues aux tableaux de contingence. Il s'agit notamment du calcul des effectifs et des fréquences. En plus nous aborderons un problème pratique qui est la notion d'indépendance entre les variables représentées dans le tableau de contingence.

Exemple de relations possibles entre les variables: taille et âge ; diabète et poids, taux de cholestérol et régime alimentaire, niche écologique et population, ensoleillement et croissance végétale, toxine et réaction métabolique, survie et pollution, effets et doses...Les caractères étudiés peuvent être aussi bien qualitatifs que quantitatifs.

III.1. TABLEAUX DE DONNEES (TABLEAU DE CONTINGENCE)

III.1.1. Définition d'une série statistique double

On considère une population d'effectif N, si on étudie deux caractères X et Y qui peuvent être qualitatives ou quantitatives de cette population, on dit que l'on étudie une série statistique double.

Les k modalités de X sont désignées par $x_1, x_2, x_3,\ldots, x_k$; $(1 \leq i \leq k)$;

Les l modalités de Y sont désignées par $y_1, y_2, y_3,\ldots, y_k$; $(1 \leq j \leq l)$;

L'ensemble des couples $(x_i ; y_j)$ définit une série statistique à deux variables.

Exemples III.1 :

- Le poids et la taille de nouveaux nés dans une maternité ;
- Le volume des ventes et le montant alloué à la publicité dans une entreprise ;
- La consommation d'un véhicule et sa vitesse.

III.1.2. Tableau de contingence

La répartition des N observations, ou distribution conjointe, suivant les modalités de X et Y se présente sous forme d'un tableau à double entrée, appelée **tableau de contingence** ou tableau à double entrée ou tableau croisé ou parfois tableau de corrélation (tableau de k lignes et de l colonnes).

Soit x la variable dont les modalités sont présentées **en ligne** et k le nombre total de ses modalités.

Soit y la variable dont les modalités sont présentées **en colonne** et l le nombre total de ses modalités.

- **i :** Indice d'une ligne ;
- $\mathbf{x_i}$ **:** Une modalité de X avec i=1,...,k ;
- **j** : Indice d'une colonne ;
- $\mathbf{y_j}$: Une modalité de Y avec j=1,...,l ;
- $\mathbf{n}_{ij}$: Effectif partiel.

Tableau III.1. Tableau de contingence.

X\Y	y_1	y_2	y_3	. . .	y_j	y_l	Totale
x_1	n_{11}	n_{12}	n_{13}	. . .	n_{1j}	n_{1l}	$n_{1\bullet}$
x_2	n_{21}	n_{22}	n_{23}	. . .	n_{2j}	n_{2l}	$n_{2\bullet}$
x_3	n_{31}	n_{22}	n_{33}	. . .	n_{3j}	n_{3l}	$n_{3\bullet}$
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
x_i	n_{i1}	n_{i2}	n_{i3}	. . .	n_{ij}	n_{il}	$n_{i\bullet}$
x_k	n_{k1}	n_{k2}	n_{k3}	. . .	n_{kj}	n_{kl}	$n_{k\bullet}$
Totale	$n_{\bullet 1}$	$n_{\bullet 2}$	$n_{\bullet 3}$	. . .	$n_{\bullet j}$	$n_{\bullet l}$	$n_{\bullet\bullet}$=N

a. L'effectif

➢ Les effectif n_{ij} désignent le nombre de fois où la modalité x_i de la variable X et la modalité y_j de la variable Y ont été observées simultanément.

➢ Les effectif $n_{i.}$ appelé effectif marginal de X, représente le nombre total des observations de la modalité x_i de X, quelle que soit la modalité de Y, on note :

$$n_{i.} = \sum_{j=1}^{l} n_{ij}$$

➢ De même les effectif $n_{.j}$ appelé effectif marginal de Y, représente le nombre total des observations de la modalité y_j de Y, quelle que soit la modalité de X, on note :

$$n_{j.} = \sum_{i=1}^{k} n_{ij}$$

On a évidemment :

$$\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij} = \sum_{j=1}^{l} n_{i.} = \sum_{i=1}^{k} n_{.j} = N$$

➢ La distribution conjointe peut aussi être définie par les fréquences :

$$f_{ij} = \frac{n_{ij}}{N}$$

Exemple III.2 :

Soit la série statistique bidimensionnelle du couple (X, Y) suivante :

Trouver les effectifs et les fréquences suivants :

n_{12} ; n_{33} ; n_{42} ; $n_{.1}$; $n_{.3}$; $n_{.4}$; $n_{2.}$; $n_{3.}$; $n_{1.}$; f_{12} ; f_{33} ; f_{42} ; $f_{.1}$; $f_{.3}$; $f_{.4}$; $f_{2.}$; $f_{3.}$; $f_{1.}$

X\Y	-2	0	2	3	$\mathbf{n_{i.}}$
2	3	4	0	6	13
3	4	3	3	2	12
4	2	3	3	2	10
$\mathbf{n_{.j}}$	9	10	6	10	35

Correction :

Les effectifs : $n_{12} = 4$; $n_{33} = 3$; n ; $_{32}$= 3$n_{.1} = 9$; $n_{.3} = 6$; $n_{.4} = 10$; $n_{2.} = 12$; $n_{3.} = 10$; $n_{1.} = 13$.

Les fréquences : $f_{12} = \frac{n_{12}}{N} = \frac{4}{35} = 0.114$; $f_{33} = 0.085$; $f_{42} = 0.085$; $f_{.1} = 0.257$; $f_{.3} = 0.171$; $f_{.4} = 0.28$; $f_{2.} = 0.342$; $f_{3.} = 0.28$; $f_{1.} = 0.37$.

III.2. REPRESENTATION GRAPHIQUE (NUAGE DE POINTS)

Il s'agit d'un graphique très commode pour représenter les observations simultanées de deux variables quantitatives.

Si les observations de deux variables statistiques X et Y sont connues individuellement, on commence par les visualiser en les représentant sous la forme d'un nuage de points : dans un repère cartésien, chaque observation (x_i ; y_i) est figurée par le point M_i de coordonnées (x_i ; y_i), et la forme du nuage donne une information sur le type d'une éventuelle liaison.

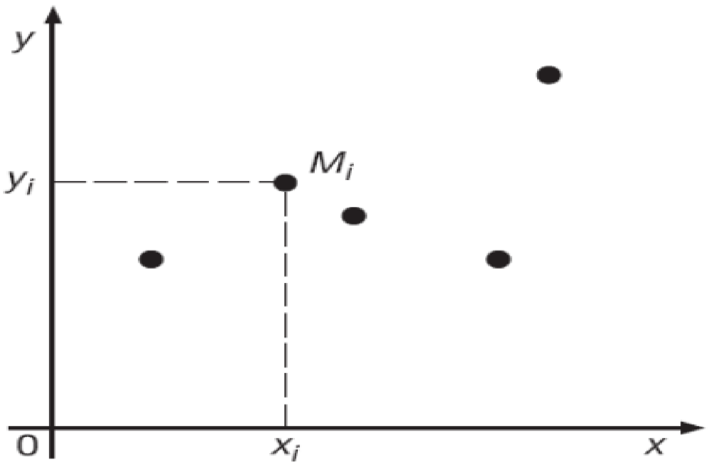


Figure III.1. Nuage de points.

III.2.1. Point moyen

III.2.1.1. Définition

Soit une série statistique à deux variables, X et Y, dont les valeurs sont des couples (x_i; y_i). On appelle **point moyen** de la série le point **G** de coordonnées :

$$\bar{\mathbf{x}} = \frac{\mathbf{x_1} + \mathbf{x_2} + \mathbf{x_3} + \cdots + \mathbf{x_n}}{\mathbf{N}}$$

$$\bar{\mathbf{y}} = \frac{\mathbf{y_1} + \mathbf{y_2} + \mathbf{y_3} + \cdots + \mathbf{y_n}}{\mathbf{N}}$$

Exemple III.3 :

Lors d'un examen d'une durée de 4 heures, on a relevé la durée de composition (c'est-à-dire au bout de combien de temps chacun a rendu sa copie) et la note (sur 20) des 12 étudiants qui se sont présentés.

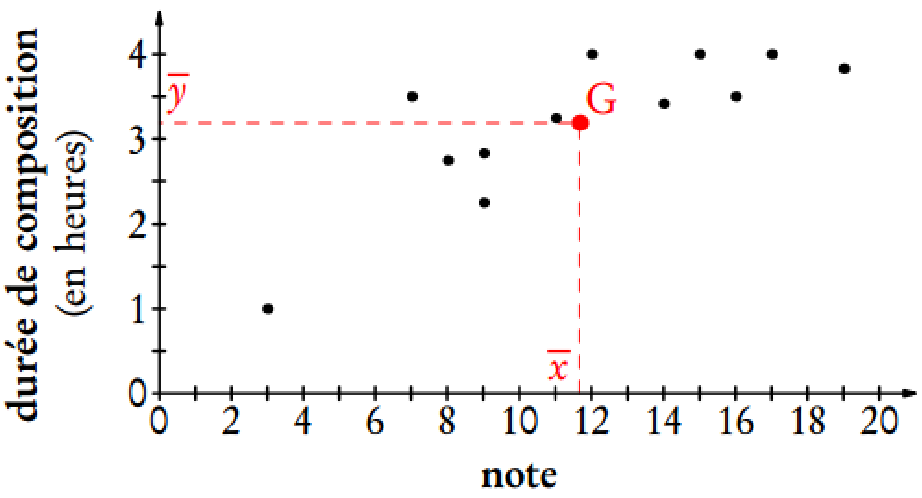
Numéro d'étudiant i	**Note (x_i)**	**Durée de composition (y_i)**
1	8	2h45
2	14	3h25
3	9	2h15
4	17	4h
5	19	3h50
6	3	1h
7	7	3h30
8	15	4h
9	12	4h
10	11	3h15
11	9	2h50
12	16	3h30

Le point moyen du nuage est le point (G) dont les coordonnées sont les moyennes marginales de la série $G = (\bar{x}; \bar{y})$:

On trouve, après calculs, les moyennes marginales suivantes :

$$\bar{x} = 11.67 \quad et \;\; \bar{y} = 3h\ 11min\ 40s;$$

On en déduit la position du point moyen sur le graphique ci-contre.



III.3. DISTRIBUTIONS MARGINALES

Disposant d'une distribution conjointe, on peut déduire les distributions marginales qui permettent d'étudier séparément chaque variable en représentant graphiquement sa distribution et s'il s'agit d'une variable quantitative, en calculant ses caractéristiques de tendance centrale et de dispersion.

On complète habituellement les tableaux de contingence par une ligne et une colonne supplémentaires dans lesquelles on calcule les sommes des colonnes et rangées respectivement.

C'est à dire sur la marge du tableau de contingence, on peut extraire les données seulement par rapport à X et seulement par rapport à Y.

Les k couples (x_i ; $n_{i.}$) forment la distribution marginale de la variable X.

Les l couples (y_j ; $n_{.j}$) forment la distribution marginale de la variable Y.

Exemple III.4 :

Reprenant la série statistique bidimensionnelle de **l'exemple III.2**.

X\Y	-2	0	2	3	**$n_{i.}$**
2	3	4	0	6	**13**
3	4	3	3	2	**12**
4	2	3	3	2	**10**
$n_{.j}$	**9**	**10**	**6**	**10**	**35**

Les distributions marginales selon X et Y sont :

a. Distribution marginale de X

X	**Effectif marginale $n_{i.}$**
2	13
3	12
4	10
Totale	35

b. Distribution marginale de Y

Y	**Effectif marginale $n_{.j}$**
-2	9
0	10
2	6
3	10
Totale	35

III.3.1. Fréquence marginale

Les distributions marginales peuvent aussi être données sous forme de fréquences :

$$f_{i.} = \frac{n_{i.}}{N} \text{ et } f_{.j} = \frac{n_{.j}}{N}$$

De plus, on a :

$$\sum_{j=1}^{l} f_{i.} = \sum_{i=1}^{k} f_{.j} = 1$$

Ces deux distributions peuvent se présenter sous forme de tableaux statistiques :

a. Distribution marginale de X

X	Effectif marginale $n_{i.}$	Fréquence marginale $f_{i.}$
x_1	$n_{1.}$	$f_{1.} = \frac{n_{1.}}{N}$
x_2	$n_{2.}$	$f_{2.} = \frac{n_{2.}}{N}$
x_3	$n_{3.}$	$f_{3.} = \frac{n_{3.}}{N}$
. . .	. . .	. . .
x_i	$n_{i.}$	$f_{i.} = \frac{n_{i.}}{N}$
x_k	$n_{k.}$	$f_{k.} = \frac{n_{k.}}{N}$
Totale	N	1

b. Distribution marginale de Y

Y	Effectif marginale $n_{.j}$	Fréquence marginale $f_{.j}$
y_1	$n_{.1}$	$f_{.1} = \frac{n_{.1}}{N}$
y_2	$n_{.2}$	$f_{.2} = \frac{n_{.2}}{N}$
y_3	$n_{.3}$	$f_{.3} = \frac{n_{.3}}{N}$
. . .	. . .	. . .
y_j	$n_{.j}$	$f_{.j} = \frac{n_{.j}}{N}$
y_l	$n_{.k}$	$f_{.l} = \frac{n_{.l}}{N}$
Totale	N	1

III.3.2. Caractéristique des séries marginales

➢ **Les moyennes marginales** des variables X et Y sont :

$\bar{x}_M = \frac{1}{N}\sum_{i=1}^{k} n_{i.}x_i = \sum_{i=1}^{k} f_{i.}x_i$ et $\bar{y}_M = \frac{1}{N}\sum_{j=1}^{l} n_{.j}y_i = \sum_{j=1}^{l} f_{.j}y_j$

➢ **Les variances marginales** des variables X et Y sont données par :

$V_M(\mathrm{x}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_{i.}}\,(\mathrm{x_i} - \bar{x}_M)^2$ et $V_M(\mathrm{y}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{j}=1}^{\mathrm{l}} \mathrm{n_{.j}}\,(\mathrm{y_j} - \bar{y}_M)^2$

➢ **Les écarts-type marginaux** de X et Y sont donnés par :

$\sigma_x = \sqrt{V_M(\mathrm{x})}$ et $\sigma_y = \sqrt{V_M(\mathrm{y})}$

Exemple III. 5 :

En reprenant l'exemple III.2 et on détermine :

- **Les moyennes marginales de X et de Y** comme suit :

$\bar{x}_M = \frac{1}{N}\sum_{i=1}^{k} n_{i.}x_i = \frac{1}{35}(2 \times 13 + 3 \times 12 + 4 \times 10) = \frac{102}{35} = 2.914.$

Et

$\bar{y}_M = \frac{1}{N}\sum_{j=1}^{l} n_{.i}y_i = \frac{1}{35}(-2 \times 9 + 0 \times 10 + 2 \times 6 + 3 \times 10) = \frac{24}{35} = 0.686.$

- **Les variances marginales des variables X et Y** sont :

$V_M(\mathrm{x}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_{i.}}\, x_i^2 - \overline{x_M}^2 = \frac{1}{35}\,(4 \times 13 + 9 \times 12 + 16 \times 10) - (2.914)^2 = 0.65.$

Et

$V_M(\mathrm{y}) = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_{.j}}\, y_j^2 - \overline{y_M}^2 = \frac{1}{35}\,(150) - (0.686)^2 = 3.815.$

- **Les écarts-type marginaux de X et Y** sont donnés par :

$\sigma_x = \sqrt{V_M(\mathrm{x})} = \sqrt{0.65} = 0.806.$

Et

$\sigma_y = \sqrt{V_M(\mathrm{y})} = \sqrt{3.815} = 1.953.$

III.4. DISTRIBUTIONS CONDITIONNELLES

La distribution de la variable Y, la variable X étant égale à x_i, est appelée distribution conditionnelle de Y pour X = x_i :

Y/X=x_i	y_1	…	y_j	…	y_l	**Totale**
Effectif	n_{i1}	…	n_{ij}	…	n_{il}	$n_{i.}$

III.4.1. Fréquence conditionnelle

Cette distribution des n_{ij} observations, satisfaisant à la condition X = x_i, est présentée sous la forme de fréquences conditionnelles :

$$\boldsymbol{f_{j/i}} = \frac{\boldsymbol{n_{ij}}}{\boldsymbol{n_{i.}}} \quad \text{avec} \quad \boldsymbol{\sum_{j=1}^{l} f_{j/i} = 1}$$

Y/X=x_i	y_1	…	y_j	…	…	y_l	**Totale**
Fréquence	$f_{1/i}$	…	$f_{j/i}$	…	…	$f_{l/i}$	1

La fréquence $f_{j/i}$ parfois notées et f_j^i se lit " f indice j si i ", c'est-à-dire fréquence de y_j si X= x_i. Il y a k distributions conditionnelles de Y pour (i = 1; … ; k):

III.4.2. Caractéristique des séries conditionnelles

Lorsque la variable Y est quantitative, on peut calculer pour chaque valeur x_i ses caractéristiques de tendance centrale et de dispersion.

➢ **La moyenne conditionnelle** :

Moyenne de X/Y=y : $\bar{x}_{Y=y} = \sum_{i=1}^{k} f_{i/j}x_i$

Moyenne de Y/X=x : $\bar{y}_{X=x} = \sum_{j=1}^{l} f_{j/i} y_j$

- **La variance conditionnelle :**

Variance de X/Y=y : $V_{Y=y} = \sum_{i=1}^{k} f_{i/j} (x_i - \bar{X}_{Y=y})^2 = \sum_{i=1}^{k} f_{i/j} x_i^2 - (\bar{x}_{Y=y})^2$

Variance de Y/X=x : $V_{X=x} = \sum_{j=1}^{l} f_{j/i} (y_i - \bar{y}_{X=x})^2 = \sum_{j=1}^{l} f_{j/i} y_j^2 - (\bar{y}_{X=x})^2$

Exemple III.6 :

En reprenant l'exemple III.2**,** alors pour déterminer la moyenne conditionnelle de X quand Y=2, il suffit d'observer le comportement de X relatif à la colonne Y=2.

X	**Y=2**
2	0
3	3
4	3
$n_{.j}$	6

$$\bar{x}_{y=2} = \frac{0 \times 2 + 3 \times 3 + 3 \times 4}{6} = 3.5$$

Pour déterminer la moyenne conditionnelle de Y quand X=3, il suffit d'observer le comportement de Y relatif à la colonne X=3 :

Y	**X=3**
-2	4
3	3
2	3
3	2
$n_{i.}$	12

$$\bar{y}_{x=3} = \frac{-2 \times 4 + 0 \times 3 + 3 \times 2 + 3 \times 2}{12} = \frac{1}{3} = 0.33$$

Remarque III.1 :

Si X et Y sont indépendantes alors : $\bar{x} = \bar{x}_{Y=y}$ et $\bar{y} = \bar{y}_{X=x}$

III.5. INDEPENDANCE DE DEUX VARIABLES

Les deux variables X et Y sont dites indépendantes si les variations de l'un des caractères n'entrainent pas de variations pour l'autre caractère. On posera alors la définition suivante :

III.5.1. Définition

Les séries statistiques $(x_i, n_{i.})$; $1 \leq i \leq k$ et $(y_j, n_{.j})$; $1 \leq j \leq l$ sont dites indépendantes, si l'on a :

$$f_{ij} = f_{i.} \times f_{.j};\ \forall\, 1 \leq i \leq k \text{ et } \forall\, 1 \leq j \leq l.$$

Remarque III.2 :

En pratique, pour montrer que deux variables ne sont pas indépendantes, il suffit de trouver un i_0 et un j_0 tels que :

$\boldsymbol{f_{i_0j_0} \neq f_{i_0.} \times f_{.j_0}}$, ce qui donne $\boldsymbol{n_{i_0j_0} \times N \neq n_{i_0.} \times n_{.j_0}}$

III.6. COVARIANCE ENTRE DEUX VARIABLES STATISTIQUES

III.6.1. Définition 1

La covariance est égale à la moyenne des écarts des couples (x_i; y_i) de X et Y par rapport au point $(\bar{x}; \bar{y})$, Elle est notée par $Cov(X, Y)$ ou S_{XY} et elle est définie par :

$$\mathbf{Cov(X, Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \bar{x}\,.\,\bar{y}}$$

III.6.2. Définition 2

Dans le cas de données groupées dans un tableau de contingence (covariance pondérée) est donnée par :

$$\mathbf{Cov(X, Y) = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}(x_i - \bar{x})\left(y_j - \bar{y}\right) = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}(x_i y_j - \bar{x}\,.\,\bar{y})}$$

La covariance indique le sens de la relation entre les variables X et Y. Ainsi, on peut distinguer les cas suivants :

- Si **Cov(X, Y)** > **0**, alors on peut dire que la relation entre les deux variables est positive. Dans ce cas, ces deux variables varient dans le même sens ;
- Si **Cov(X, Y)** < **0**, alors on peut dire que la relation entre les deux variables est négative. Dans ce cas, ces deux variables varient en sens inverse ;
- Si **Cov(X, Y) = 0**, alors on peut dire qu'il n'y a pas de relation entre les deux variables. Dans ce cas, les variations de l'une n'entrainent pas la variation de l'autre.

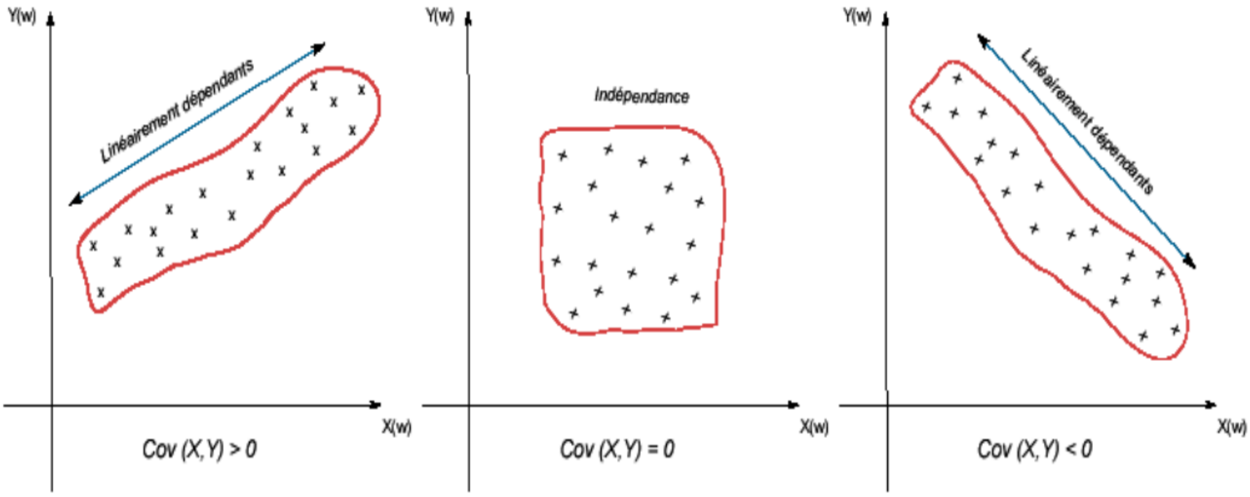


Figure III.2. La Covariance et la variabilité.

III.6.3. Propriétés de la covariance

- $Cov(X,Y) = Cov(Y,X)$;
- $Cov(X,X) = V(x)$;
- $Cov(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$;
- $\forall\, a, b, x_0, y_0 \in \mathbb{R}$: $Cov(aX + x_0 bY + y_0) = ab\ Cov(X,Y)$;
- $\mathrm{V(aX + bY + c)} = a^2\, V(X) + b^2 V(Y) + 2abCov(X,Y)$;
- $|Cov(X,Y)| \leq \sqrt{V(X).V(Y)}$;

III.7. COEFFICIENT DE CORRELATION LINEAIRE

Nous allons calculer le coefficient de corrélation entre deux séries de même longueur. On suppose qu'on a les tableaux de valeurs suivants : X(x_1;…; x_k) et Y (y_1;…; y_l) pour chacune des deux séries.

III.7.1. Définition

C'est un indice qui mesure le degré de liaison entre deux variables X et Y. On appelle coefficient de corrélation linéaire ou coefficient de Bravais Pearson entre deux variables statistiques X et Y, le rapport de leur covariance par le produit de leurs écarts-types, il est noté par $\mathbf{r}(\boldsymbol{X}, \boldsymbol{Y})$ ou $\boldsymbol{\rho_{XY}}$ et il est défini par :

$$\mathbf{r(X,Y)} = \frac{\mathbf{Cov(X,Y)}}{\boldsymbol{\sigma_x \sigma_y}}$$

La liaison entre deux variables numériques peut être étudiée grâce au coefficient de corrélation. Néanmoins, il faut bien garder présent à l'esprit que le coefficient de corrélation de Bravais-Pearson ne mesure que des relations linéaires, et sa valeur n'est en rien le reflet de l'existence d'un lien de causalité entre les deux variables.

III.7.2. Propriétés du coefficient de corrélation linéaire

Le coefficient de corrélation est toujours compris entre -1 et +1 :

- Si r = +1 alors les points se trouvent tous sur une même droite croissante, la corrélation linéaire positive parfaite ;
- Si r = -1 alors les points se trouvent tous sur une même droite décroissante, la corrélation linéaire négative parfaite ;
- Si r = 0 alors il n'y a pas une relation linéaire entre les variables X et Y.

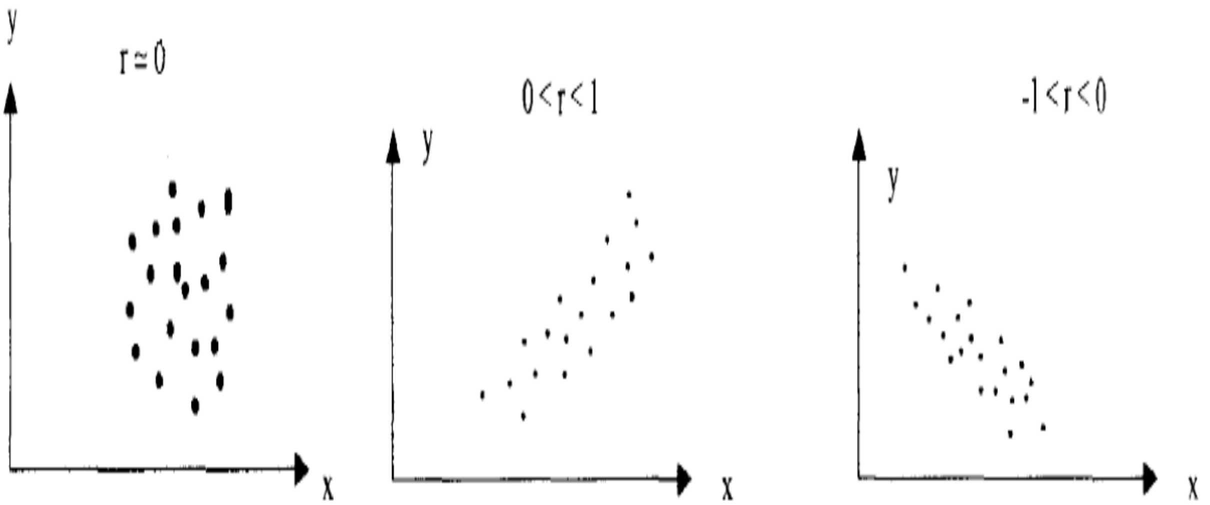


Figure III.3. Exemples de nuages de points et coefficients de corrélation.

III.8. AJUSTEMENT

Il est très intéressant pour l'analyse d'une série de données de déterminer une fonction dont le graphe approche au maximum le nuage de points. C'est cette approximation que l'on appelle ajustement statistique.

Dans certains cas, le nuage de points laisse prévoir une relation fonctionnelle globale entre X et Y mais cette relation n'est pas nécessairement affine (régression linéaire). Nous sommes alors amenés à réaliser un ajustement se ramenant, par un changement de variable simple, à un ajustement affine.

III.8.1. Interpolation, extrapolation

L'objectif de l'ajustement est, à partir des valeurs d'une série statistique à deux variables, d'obtenir des approximations pour des valeurs inconnues de cette série.

III.8.1.1. Définitions

L'interpolation et l'extrapolation sont des méthodes qui consistent à estimer une valeur inconnue dans une série statistique.

- Pour **une interpolation**, le calcul est réalisé dans le domaine d'étude fourni par les valeurs de la série.
- Pour **une extrapolation**, le calcul est réalisé en dehors du domaine d'étude.

Exemples III.8 :

- On donne une série exprimant la population d'une ville en fonction des années et on souhaite faire des prévisions pour les années à venir.

Les prévisions sortent du domaine d'étude de la série, on parle dans ce cas **d'extrapolation**.

- On donne une série exprimant la température extérieure et la consommation électrique correspondante. Les températures étudiées s'échelonnent entre -10°C et 10°C avec un pas de 4°C. Sans faire de nouveaux relevés, on souhaite estimer la consommation électrique pour toutes les températures entières comprises entre -10°C et 10°C.

Les calculs sont dans le domaine d'étude de la série, on parle dans ce cas **d'interpolation**.

III.8.2. Ajustement linéaire

III.8.2.1. Droite de régression

Dans le cas particulier où l'on a pu mettre en évidence l'existence d'une relation linéaire significative entre deux caractères X et Y, on peut chercher à formaliser la relation moyenne qui unit ces deux variables à l'aide de trois équations, ces trois équations correspondent à trois droites différentes, trois résumés différents du nuage de points $M_i(x_i, y_i)$. La différence entre les trois droites vient du fait que les trois équations proposées correspondent à trois objectifs différents :

a. a.X + b.Y + c = 0 : Equation de la droite moyenne liant les caractères X et Y

Cette droite moyenne est un résumé de la relation entre X et Y qui n'introduit aucune hypothèse particulière sur le sens de la dépendance causale qu'il peut y avoir entre les deux variables. Elle visera donc à tracer la droite qui soit la plus proche de tous les points, c'est-à-dire les résidus définis par la perpendiculaire de chaque point à la droite moyenne (plus court chemin).

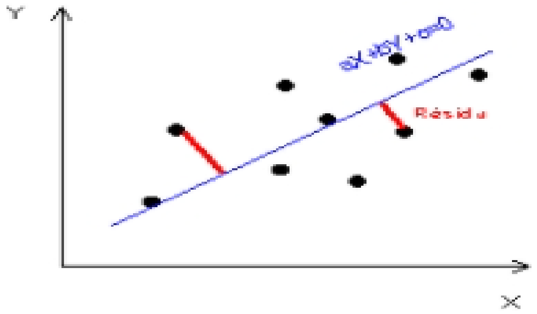
b. Y = a.X + b : Droite de régression de Y en fonction de X

Cette droite de régression de Y en fonction de X introduit l'hypothèse que les valeurs de Y dépendent de celles de X, c'est-à-dire postulent que la connaissance des valeurs de X permet de prévoir les valeurs de Y. Il s'agit donc d'un modèle de prévision et l'objectif est de minimiser l'erreur de prévision c'est-à-dire la distance entre les valeurs Y_i observées et les valeurs Y^*i estimés par la relation $Y^*=aX+b$. Les résidus seront donc la distance à la droite par rapport à l'axe Oy.

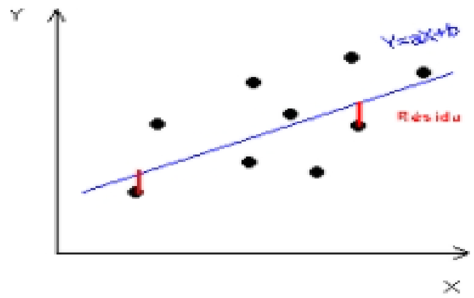
c. X = a.Y + b : Droite de régression de X en fonction de Y

Cette droite de régression de X en fonction de Y introduit l'hypothèse inverse que les valeurs de X dépendent de celles de Y, c'est-à-dire postulent que la connaissance des valeurs de Y permet de prévoir les valeurs de X. Il s'agit donc cette fois-ci de minimiser l'erreur de prévision sur X c'est-à-dire la distance entre les valeurs X_i observées et les valeurs X^*_i estimés par la relation $X^*=aY+b$. Les résidus seront donc la distance à la droite par rapport à l'axe Ox.

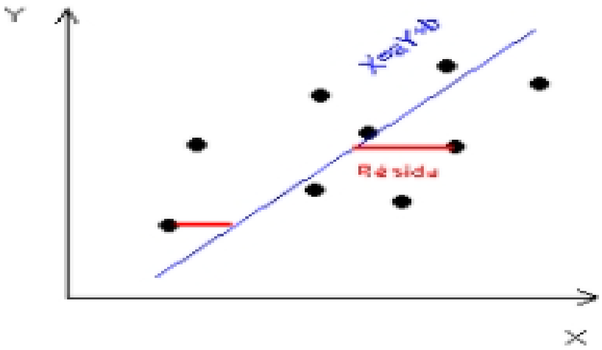
La figure III.4 suivante représente les trois manières différentes de résumer un nuage de point.



1. Droite exprimant la relation moyenne entre X et Y, $aX + bY + c = 0$



2. Droite exprimant Y en fonction de X, $Y = aX + b$



3. Droite exprimant X en fonction de Y, $X = aY + b$

Figure III.4. Trois manières différentes de résumer un nuage de point.

Comme on peut le voir sur la figure III.4, les droites de régression linéaire obtenues seront différentes en fonction de l'hypothèse faite sur la relation entre X et Y et la présence ou non d'une dépendance entre les deux caractères. Il convient donc de toujours spécifier l'hypothèse qui est faite avant d'entreprendre le calcul d'une droite de régression.

III.8.2.2. Méthode des points moyens (ou méthode de Mayer)

La méthode des points moyens propose d'ajuster le nuage de points par une droite passant par les deux points moyens G_1 et G_2 de deux ensembles de points du nuage, l'un formé des points les plus à gauche, et l'autre formé des points les plus à droite. Ainsi, ces deux ensembles forment une partition du nuage de points et contiennent le même nombre de points (plus un pour l'un si n est impair).

Ainsi, pour G_1 de coordonnées $(\bar{x}_1; \bar{y}_1)$ et G_2 de coordonnées$(\bar{x}_2; \bar{y}_2)$, la méthode des points moyens propose la droite d'équation y = a + bx, avec :

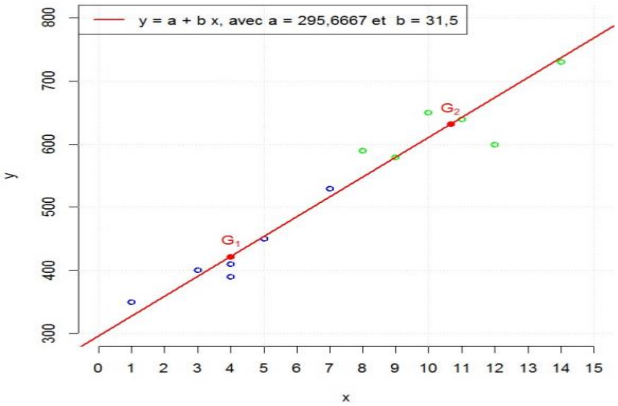
$$\boldsymbol{b = \frac{\bar{y}_2 - \bar{y}_1}{\bar{x}_2 - \bar{x}_1}\ ;\ a = \bar{y}_1 - b\bar{x}_1}$$

Exemples III.9 :

Une étude a été menée auprès de 12 étudiants afin d'expliquer le score à un examen de mathématiques à partir du temps consacré à la préparation de cet examen. Pour chaque étudiant, on dispose du temps de révision en heures (variable X) et du score obtenu sur 800 points (variable Y). Les résultats sont :

x_i	4	9	10	14	4	7	12	1	3	8	11	5
y_i	390	580	650	730	410	530	600	350	400	590	640	450

La méthode des points moyens propose la droite suivante :



On a alors considéré deux ensembles de points du nuage. L'un est formé des points les plus à gauche (en bleue) :

M_8	M_9	M_1	M_5	M_{12}	M_6
(1 ; 350)	(3 ; 400)	(4 ; 390)	(4 ; 410)	(5 ; 450)	(7 ; 530)

L'autre est formé des points les plus à droite (en vert) :

M_{10}	M_2	M_3	M_{11}	M_7	M_4
(8 ; 590)	(9 ; 580)	(10 ; 650)	(11 ; 640)	(12 ; 600)	(14 ; 730)

On a déterminé les points moyens G_1 et G_2 de ces ensembles : Ainsi, **G_1 est de coordonnées** :

$$(\bar{x}_1; \bar{y}_1) = (\frac{1+3+4+4+5+7}{6}; \frac{350+400+390+410+450+350}{6}) = (4; 421.66)$$

Et **G_2 est de coordonnées** :

$$(\bar{x}_2; \bar{y}_2) = (\frac{8+9+10+11+12+14}{6}; \frac{590+580+650+640+600+730}{6}$$

$$= (10.66; 631.66)$$

En utilisant ces coordonnées, l'équation de la droite passant par G_1 et G_2 est **y = a + bx,** avec :

$b = \frac{631.66-421.66}{10.66-4} = 31.5$, $a = 421.66 - 31.5 \times 4 = 295.66$.

Avec cette équation, on peut alors faire des prévisions. Par exemple, une valeur estimée du score d'un étudiant ayant consacré 16 heures de préparation à l'examen est :

$$\boldsymbol{y = a + bx = 295,66 + 31,5 \times 16 = 799,66.}$$

Ainsi, on prévoit un score de 800 pour un tel étudiant.

III.8.2.3. Méthode des moindres carrées

L'idée est de transformer un nuage de point en une droite. Celle-ci doit être la plus proche possible de chacun des points.

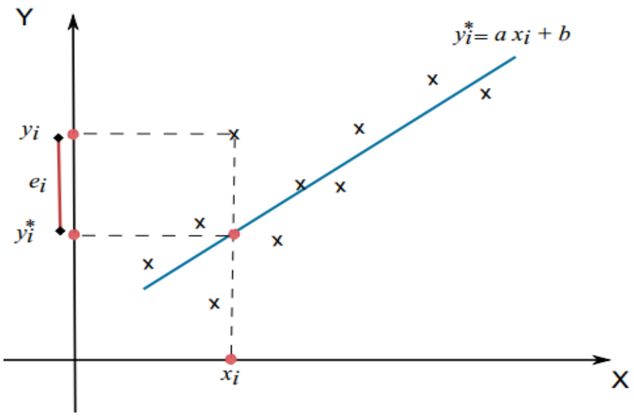


Figure III.5. La droite la plus proche possible de chacun des points.

On cherchera donc à minimiser les écarts entre les points et la droite. Pour cela, on utilise la méthode des moindres carrées. Cette méthode vise à expliquer un nuage de points par une droite qui lie Y à X, c'est-à-dire :

$$\boldsymbol{y = ax + b}$$

Telle que la distance entre le nuage de points et droite soit minimale. Cette distance matérialise l'erreur, c'est à dire la différence entre le point réellement observé et le point prédit par la droite. Si la droite passe au milieu des points, cette erreur sera alternativement positive et négative, la somme des erreurs étant par définition nulle. Ainsi, la méthode des moindres carrés consiste à chercher la valeur des paramètres a et b qui minimise la somme des erreurs élevées au carré.

Donc, la droite de régression, qui rend la distance entre elle et les points minimale, est donnée par :

$$\boldsymbol{D(Y/X): Y = aX + b}$$

Avec :

$$\boldsymbol{a = \frac{Cov(X,Y)}{V(X)}} \quad \text{et} \quad \boldsymbol{b = \bar{y} - a\bar{x}}$$

Ou bien :

$$\boldsymbol{D(X/Y): X = a'Y + b'}$$

Avec :

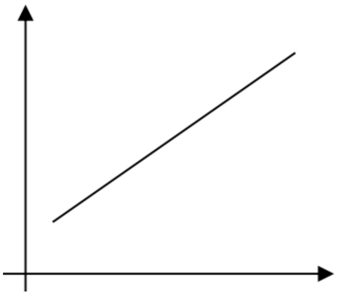
$$\boldsymbol{a' = \frac{Cov(X,Y)}{V(Y)}} \quad \text{et} \quad \boldsymbol{b' = \bar{x} - a\bar{y}}$$

Remarques III.3 :

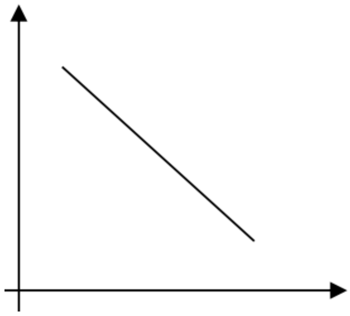
Les deux droites d'estimation sont différentes, mais on ne peut dire laquelle qui représente un meilleur ajustement.

Le coefficient de corrélation linéaire est égal au produit des pentes : $(r(X,Y))^2 = aa'$

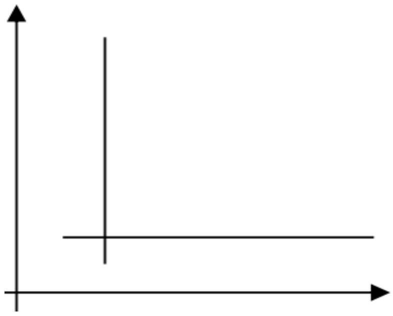
- Pour s'assurer que l'ajustement est valide, on calcule le coefficient de corrélation, et s'il est voisin en valeur absolue de 1, l'ajustement est valide, $(0.7 < |r(X,Y)| < 1)$.
- Si $|r(X,Y)| = 1$ alors les points sont alignés ;
- Si $r(X,Y) = 0$ alors X et Y sont non corrélées ;
- Si $r(X,Y) > 0$ alors X et Y croient dans le même sens ;
- Si $r(X,Y) < 0$ alors X et Y croient dans le sens différent.



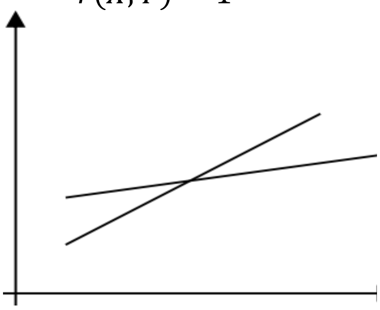
$r(X,Y) = 1$



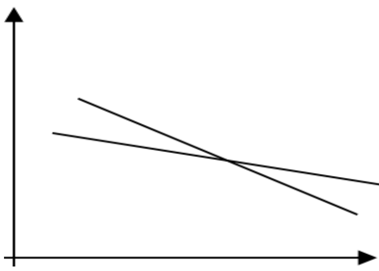
$r(X,Y) = -1$



$r(X,Y) = 0$



$r(X,Y) > 0$



$r(X,Y) < 0$

Prévision :

La droite de régression de y en x, D(Y/X) permet de prédire une valeur y pour une valeur x_0 donnée : $\boldsymbol{y = ax_0 + b}$

La droite de régression de x en y (y) permet de prédire une valeur x pour une valeur y_0 donnée : $\boldsymbol{x = a'y_0 + b'}$

III.8.3. Ajustement non linéaire

Dans certains cas, l'ajustement à une fonction linéaire n'est pas adéquat : un ajustement des données à une fonction non linéaire doit être envisagé. Les deux cas que nous considérons sont ceux où on peut se ramener par simple transformation à un ajustement affine.

a. Ajustement à une fonction puissance

Supposons que les variables statistiques X et Y sont liées par une relation de la forme : $Y = bX^a$.

Dans ce cas, cette équation peut être transformée en prenant le logarithme :

$\ln Y = a \ln X + \ln b$.

En effectuant les changements de variables suivants :$V = \ln Y$, $U = \ln X$, $B = \ln b$, nous nous ramenons au cas étudié $V = aU + B$.

b. Ajustement à une fonction exponentielle

Supposons que les variables statistiques X et Y sont liées par une relation de la forme :

$$\mathrm{Y} = \mathrm{b}\, e^{aX}.$$

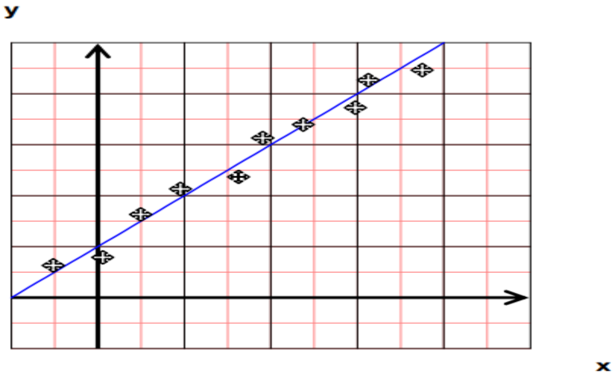
Dans ce cas, cette équation peut être transformée en passant aux logarithmes :

$$\ln Y = aX + \ln b$$

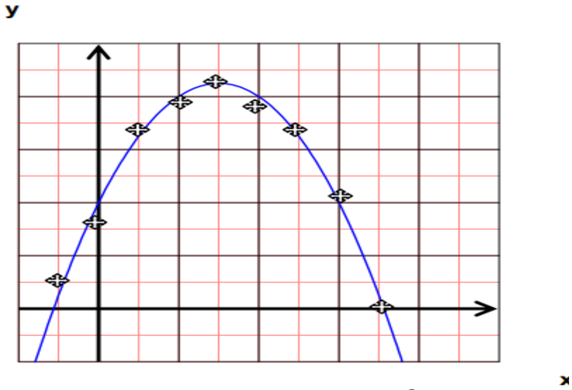
En effectuant les changements de variables suivants : $V = \ln Y$, $B = \ln b$, nous nous ramenons au cas étudié $V = a\,X + B$.

III.8.4. Récapitulatifs

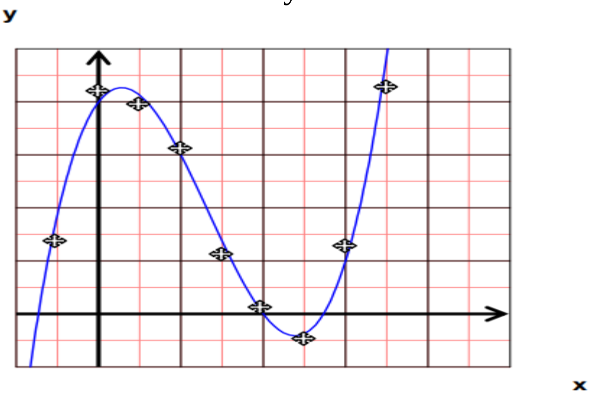
L'observation des points tracés par un outil graphique (calculatrice ou logiciel) nous permet de déterminer le modèle d'ajustement le plus approprié. Ce modèle est choisi en observant la forme du nuage de points parmi les ajustements suivants :



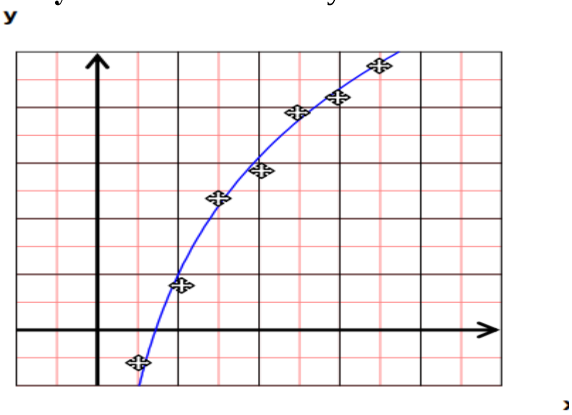
Affine : $y = ax + b$



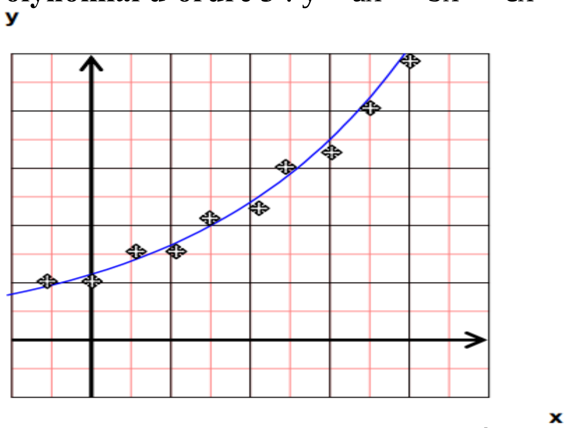
Polynomial d'ordre 2 : $y = ax^2 + bx + c$



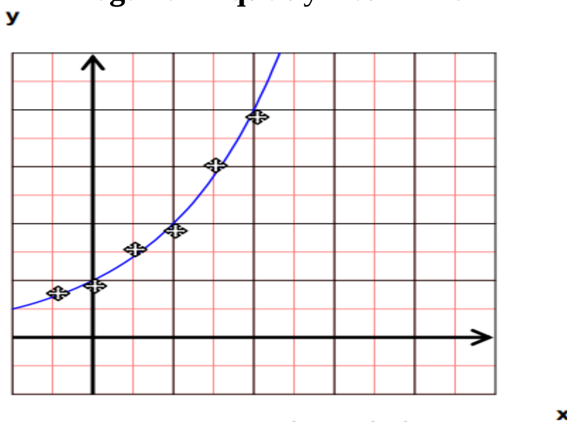
Polynomial d'ordre 3 : $y = ax^3 + bx^2 + cx + d$



Logarithmique : $y = a.\ln x + b$



Exponentiel : $y = a.b^x$ ou $y = a.e^{bx}$



Puissance : $y = a^x$

Partie B : Probabilités

Dans des domaines très différents comme les domaines scientifiques, sociologique ou médical, on s'intéresse à de nombreux phénomènes dans lesquels apparaît l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre.

Une expérience est appelée "aléatoire" s'il est impossible de prévoir à l'avance son résultat et si, répétée dans des conditions identiques, elle peut donner des résultats différents :

- *Succession d'appels à un standard téléphonique non surchargé ;*
- *Observation de la durée de vie d'un individu anonyme dans une population ;*
- *Observation de la durée de fonctionnement sans panne d'appareil ;*
- *Jeu de pile ou face.*

Voici d'autres exemples de domaines d'applications des probabilités.

***Fiabilité** : On considère un système formé par plusieurs composants. On s'intéresse à la fiabilité du système : on va chercher à calculer la probabilité que le système fonctionne encore à un instant donné. Il faut pour cela connaître la probabilité que chacun des composants fonctionne à cet instant et tenir compte du fait que les composants ne fonctionnent peut-être pas indépendamment les uns des autres.*

***Fatigue des matériaux** : Les données de fatigue des matériaux sont très dispersées. On fait alors appel à des modélisations probabilistes et à des méthodes statistiques afin, par exemple, de construire des intervalles de confiance pour le nombre moyen de cycles jusqu'à la rupture.*

***Télécommunications :** En télécommunications, on doit souvent tenir compte du "bruit" dans les systèmes. Par exemple, supposons qu'un système émet soit un 0, soit un 1, et qu'il y a un risque p que le chiffre émis soit mal reçu. Il est alors intéressant de calculer la probabilité qu'un 0 ait été émis, sachant qu'un 0 a été reçu, ou encore la probabilité qu'il y ait une erreur de transmission.*

CHAPITRE I. ANALYSE COMBINATOIRE

L'analyse combinatoire est une branche des mathématiques qui étudie comment compté les objets. Elle fournit des méthodes de dénombrements particulièrement utiles en théorie des probabilités. Les probabilités dites combinatoires utilisent constamment les formules de l'analyse combinatoire développées dans ce chapitre. Un exemple des applications intéressantes de cette dernière est la démonstration du développement du binôme de Newton utilisé dans le calcul des probabilités d'une loi binomiale.

I.1. PRINCIPE FONDAMENTAL DE DENOMBREMENT

Ce principe de dénombrement sera essentiel par la suite, il établit en gros que si une expérience peut produire **m** résultats et une autre **n**, alors il y a **mn** résultats possibles lorsqu'on considère ces deux expériences ensemble.

Théorème :

Supposons qu'il y a deux expériences à réaliser. Si l'expérience **1** peut produire l'un quelconque de **m** résultats et si, pour chacun d'entre eux, il y a **n** résultats possibles pour l'expérience **2**, alors il existe **mn** résultats pour les deux expériences prises ensemble.

Principe fondamental généralise

Lorsqu'il y a plus de deux expériences à réaliser, le principe fondamental de dénombrement peut-être généraliser comme suit :

Si r expériences doivent être réalisées, et s'il y a ni résultats possibles pour la $i^{éme}$ expérience, $i = 1, \ldots, r$, il aura alors un total de $n_1 \times n_2 \times n_3 \times \ldots \times n_r$ résultats pour les r expériences prises ensemble.

Exemple I.1 :

Une petite communauté se compose de dix hommes et de leurs fils, chaque homme ayant trois fils. Si un homme et l'un de ses fils doivent être désigner « père et fils exemplaire », combien y a-t-il de choix différents possibles ?

Solution :

En considérant le choix du père comme la première expérience et ensuite le choix de l'un de ses fils comme la second expérience, nous concluons d'après le principe fondamental de dénombrement qu'il y a $10 \times 3 = 30$ choix possibles.

I.2. ARRANGEMENTS

I.2.1. Définition

Etant donné un ensemble E de n objets, on appelle **arrangements** de p objets toutes **suites ordonnées** de p objets pris parmi les n objets. Le nombre d'arrangements de p objets pris parmi n est noté ($\boldsymbol{A_n^p}$).

Remarque I.1 :

On a nécessairement: $1 \leq p \leq n \; et \; n, p \; \in N^*$, si $n < p$, alors $A_n^p = 0$.

Deux arrangements de p objets sont donc **distincts** s'ils diffèrent par la nature des objets qui les composent ou par leur ordre dans la suite.

Exemples I.2 :

- Une séquence d'ADN est constituée d'un enchaînement de 4 nucléotides [A (Adénine), C (Cytosine), G (Guanine) et T (Thymine)]. Il existe différents arrangements possibles de deux nucléotides ou dinucléotides avec **p=2** et **n=4**.
- Le nombre de mots de 5 lettres (avec ou sans signification) formés avec les 26 lettres de l'alphabet correspond au nombre d'arrangements possibles avec **p=5** et **n=26**.
- Le tiercé dans l'ordre lors d'une course de 20 chevaux constitue un des arrangements possibles avec **p=3** et **n=20**.

Dans les exemples précédents, l'ordre des éléments dans la suite est essentiel. Ainsi pour le deuxième exemple, le mot NICHE est différent du mot CHIEN.

Mais dans les deux premiers exemples, une base ou une lettre de l'alphabet **peut se retrouver plusieurs fois** alors que dans le troisième exemple, les trois chevaux à l'arrivée sont forcément **différents**. Il faut donc distinguer le nombre **d'arrangements avec répétition** et le nombre **d'arrangements sans répétition** (arrangements au sens strict).

I.2.2. Arrangements avec répétitions

Lorsqu'un objet peut être observé **plusieurs fois** dans un arrangement, le nombre **d'arrangement avec répétition** de p objets pris parmi n, est alors :

$$\boldsymbol{A_n^p = n^p} \quad \text{avec} \quad \boldsymbol{1 \leq p \leq n}$$

Démonstration :

- Pour le premier objet tiré, il existe n manières de ranger l'objet parmi n.
- Pour le second objet tiré, il existe également n possibilités d'arrangements car le premier objet fait de nouveau parti des n objets. On parle de **tirage avec remise**.

➢ Ainsi pour les p objets tirés, il y aura $n\ x\ n\ x\ n\ x.....x\ n$ (p fois) arrangements possibles, soit :

$$\boldsymbol{A_n^p = n^p = n \times n \times n \times \ldots \times n = n^p}$$

Exemples I.3 :

Concernant l'exemple de **la séquence d'ADN**, le nombre de dinucléotides attendus si l'on fait l'hypothèse qu'une base peut être observée plusieurs fois dans la séquence (ce qui correspond effectivement à la réalité) est donc : $A_4^2 = 4^2 = 4 \times 4 = \mathbf{16}$;

Donc on aura **16 dinucléotides possibles**.

Les 16 dinucléotides identifiables dans une séquence d'ADN sont :

AA	AC	AG	AT	CA	**CC**	CG	CT
GA	GC	**GG**	GT	TA	TC	TG	**TT**

I.2.3. Arrangements sans répétition

Lorsque chaque objet ne peut être observé qu'**une seule** fois dans un arrangement, le nombre d'**arrangements sans répétition** de p objets pris parmi n est alors :

$\boldsymbol{A_n^p} = \frac{\boldsymbol{n}!}{(\boldsymbol{n-p})!}$ avec $\boldsymbol{1 \leq p \leq n}$

Démonstration :

➢ Pour le premier objet tiré, il y a n manières de ranger l'objet parmi n.

➢ Pour le second objet tiré, il n'existe plus que n-1 manières de ranger l'objet car le premier objet ne peut plus être pris en compte. On parle de **tirage sans remise**.

➢ Ainsi pour les p objets tirés parmi n, si $1 \leq p \leq n$, il y aura :

$A_n^p = n(n-1)(n-2)(n-3) \ldots (n-p+1)$ (p produits)

De plus : $A_n^p = n(n-1)(n-2)(n-3) \ldots (n-p+1) \times \frac{(n-p)\times\ldots\times2\times1}{(n-p)\times\ldots\times2\times1}$

D'où : $\boldsymbol{A_n^p} = \frac{\boldsymbol{n}!}{(\boldsymbol{n-p})!}$

Rappel :

Si $n \in N^*$ on appelle **factorielle *n***, notée $n!$, le produit des ***n*** premiers entiers :

➢ $n! = 1 \times 2 \times 3 \times \ldots p \times (p+1) \times \ldots \times (n-1) \times n$

➢ $0! = 1$, par convention car $0! = 1$ n'est en principe pas définie.

➢ Dès que n dépasse la dizaine, **n!** se compte en millions. Il est bon de connaître la formule d'approximation suivante (« **formule de Stirling** ») : $\boldsymbol{n! \approx (\frac{n}{e})^n \sqrt{2\pi n}}$.

Exemple I.4 :

Concernant l'exemple de la séquence d'ADN, le nombre de dinucléotides attendu dans une séquence si l'on fait l'hypothèse qu'une base n'est observée qu'une seule fois est donc :

$A_4^2 = \frac{4!}{(4-2)!} = \mathbf{12}$

Sous cette contrainte, les 12 dinucléotides possibles sont :

~~AA~~	AC	AG	AT	CA	~~CC~~	CG	CT
GA	GC	~~GG~~	GT	TA	TC	TG	~~TT~~

Ceci correspond aux 16 arrangements possibles avec répétition ($\boldsymbol{A_n^p = n^p}$) auxquels sont soustraits les 4 dinucléotides (n) résultant de l'association d'une même base.

I.3. PERMUTATION

I.3.1. Permutations sans répétition

Etant donné un ensemble E de n objets, on appelle **permutations** de n objets distincts toutes **suites ordonnées** de n objets ou tout **arrangement *n* à *n*** de ces objets.

Le nombre de permutations de n objets est noté : $\boldsymbol{P_n = n!}$

La permutation de n objets constitue un cas particulier d'arrangement sans répétition de p objets pris parmi n lorsque $p = n$.

Ainsi le nombre de permutations de n objets est : $\boldsymbol{A_n^n = \frac{n!}{(n-n)!} = n!}$

Exemple I.5 :

Le nombre de manières de placer 8 convives autour d'une table est :

$P_8 = 8! = \mathbf{40320}$ **Possibilités**.

I.3.1. Permutations avec répétitions

Dans le cas où il existerait plusieurs répétitions k d'un même objet parmi les n objets, le nombre de permutations possibles des n objets doit être rapporté aux nombres de permutations des k objets identiques.

Le nombre de permutations de n objets est alors : $\mathbf{P_n = \frac{n!}{k!}}$

En effet, les permutations de k objets identiques sont toutes identiques et ne comptent que pour une seule permutation.

Exemple I.6 :

Considérons le mot « CELLULE ». Le nombre de mots possibles (avec ou sans signification) que l'on peut écrire en permutant ces 7 lettres est :

$P_7 = \frac{7!}{2!3!} = \mathbf{420}$ **Mots possibles**

En considérant deux groupes de lettres identiques : L (3 fois) et E (2 fois).

I.4. COMBINAISONS

I.4.1. Définition

Si l'on reprend **l'exemple de la séquence d'ADN**, à la différence des arrangements où les dinucléotides AC et CA formaient deux arrangements distincts, ces derniers ne formeront qu'une seule combinaison. Pour les combinaisons, on ne parle plus de **suite** ni de **série** puisque la **notion d'ordre des objets n'est plus prise en compte**. On parle alors de **tirages avec ou sans remise.**

I.4.2. Combinaisons sans remise

Etant donné un ensemble *E* de *n* objets, on appelle combinaisons de *p* objets tout ensemble de *p* objets pris parmi les *n* objets sans remise.

Le nombre de combinaisons de *p* objets pris parmi *n* est noté :$\boldsymbol{C_n^p}$.

Remarque I.2 :

On a nécessairement $1 \leq p \leq n \; et \; n, p \; \in N^*$, si $n < p$, alors $C_n^p = 0$

Exemples I.7 :

- Le tirage au hasard de 5 cartes dans un jeu de *32* (main de poker) est une combinaison avec *p=5* et *n=32*.
- La formation d'une délégation de 5 personnes parmi un groupe de 50 constitue une combinaison avec *p=5* et *n=50*.

Pour ces deux exemples, les objets tirés sont clairement distincts.

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

$$\boldsymbol{C_n^p = \frac{n!}{p!\,(n-p)!}} \; notée \; \boldsymbol{\binom{n}{p}} \; avec \; \boldsymbol{1 \leq p \leq n}$$

Démonstration :

- Pour calculer ce nombre, on utilise le principe de la division ;
- Il y A_n^p manières de tirer *p* objets parmi *n* en **les ordonnant** soit $\boldsymbol{A_n^p = \frac{n!}{(n-p)!}}$;
- Une fois les *p* objets tirés, il y a p! **manières de les ordonner** ;
- Il y a donc $\frac{A_n^p}{p!}$ manières de tirer *p* objets parmi *n* **sans** les ordonner ;

$$\boldsymbol{C_n^p = \frac{A_n^p}{p!} = \frac{1}{p!}\frac{n!}{(n-p)!}}$$

Remarque I.3 :

A la notation ancienne C_n^p, on préfère parfois la notation moderne $\binom{n}{p}$. Les nombres *n* et *p* constituent **les coefficients binomiaux**.

Exemples I.8 :

Dans le cadre de l'exemple de **la séquence d'ADN**, le nombre de dinucléotides attendus sans tenir compte de l'ordre des bases dans la séquence (hypothèse justifiée dans le cas de l'ADN non codant) est donc :

$C_n^p = \binom{n}{p} = \frac{4!}{2!(4-2)!} = \frac{4\times 3}{2\times 1} =$ **6 Dinucléotides.**

Les 6 dinucléotides possibles sous cette hypothèse sont :

AC	AG	AT	CG	CT	GT
CA	GA	TA	GC	TC	TG

Ceci correspond aux 12 arrangements possibles sans répétitions ($A_n^p = \frac{n!}{(n-p)!}$) divisé par le nombre de permutations possibles avec 2 nucléotides ($P_p = p!$).

I.4.3. Combinaisons avec remise

Le nombre de **combinaisons** de *p* objet parmi *n* **avec remise** est :

$$\boldsymbol{A_{n+p-1}^p = \frac{(n+p-1)!}{p!(n-1)!}}$$

Démonstration :

Soit la constitution de mots de 3 lettres à partir d'un alphabet à 5 lettres avec remise, on distingue 3 cas possibles :

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

D'où au total : $C_5^3 + 2C_5^3 + C_5^1 = C_7^3$ en utilisant la formule des **combinaisons** composées ou formule de Pascal.

En effet $C_5^3 + C_5^2 = C_6^3$ et $C_5^2 + C_5^1 = C_6^2$ d'où $C_6^3 + C_6^2 = C_7^3$ soit $C_7^3 = 35$ mots possibles de 3 lettres à partir d'un alphabet à 5 lettres.

Ainsi : $C_7^3 = C_{5+3-1}^3 = C_{n+p-1}^p$ avec $n = 5$ *et* $p = 3$.

I.4.4. Propriétés des combinaisons

I.4.4.1. La symétrie

Le nombre de combinaisons de *p* objets pris parmi *n* étant : $C_n^p = \frac{1}{p!}\frac{n!}{(n-p)!}$, alors :

- $C_n^0 = C_n^n = 1$ car $C_n^0 = C_n^n = \frac{n!}{n!}$
- Si $n \geq 1$ $C_n^1 = C_n^{n-1} = 1$ avec $C_n^1 = C_n^{n-1} = \frac{n!}{n-1!}$
- Si $n \geq 2$ $C_n^2 = C_n^{n-2} = \frac{n(n-1)}{2}$ avec $C_n^2 = C_n^{n-2} = \frac{n!}{2!n-2!} = \frac{n(n-1)(n-2)!}{2!n-2!}$

Par récurrence, on déduit des relations précédentes, la propriété de **symétrie** à savoir :

Si $\boldsymbol{0 \leq p \leq n,\ C_n^{n-p} = \frac{n!}{p!(n-p)!}}$ ainsi $\boldsymbol{C_n^p = C_n^{n-p}}$

Il revient au même de donner la combinaison des ***p* objets choisis** ou bien celle des ***(n-p)* qui ne le sont pas**.

I.4.4.2. Combinaisons composées ou Formule de Pascal

Si $0 \leq p \leq n-1$ $C_{n-1}^{p-1} + C_{n-1}^{p} = C_n^p$

Démonstration :

Parmi les n objets, on considère un objet en particulier.

- Si cet objet fait partie des *p* objets tirés, il y a C_{n-1}^{p-1} possibilités de choisir les *p-1* autres objets parmi les *n-1* objets restants.
- Si en revanche, l'objet ne fait pas partie du tirage, il y a C_{n-1}^{p-1} possibilités de choisir les *p* autres objets parmi les *n-1* objets restants.

D'où la relation : $C_{n-1}^{p-1} + C_{n-1}^{p-1} = C_n^p$

Les termes du **triangle de Pascal** résultent de l'application directe de cette relation.

		p-1	p	
n-1		C_{n-1}^{p-1}	C_{n-1}^{p}	
n			C_n^p	

Pour établir le triangle de Pascal, il suffit de porter les valeurs prises par p en colonne et celles prises par n en ligne (voir tableau ci-dessus). La valeur attribuée à chaque case, C_n^p, est obtenue en faisant la somme de la valeur de la case située juste au-dessus, C_{n-1}^{p} et la valeur de la case située au-dessus et à gauche C_{n-1}^{p-1} Ceci correspond à l'application de la propriété énoncée précédemment.

Le triangle de Pascal permet d'obtenir par récurrence les coefficients numériques ou **coefficient binomiaux** du binôme de Newton.

I.5. FORMULE DU BINOME DE NEWTON

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

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

Elever (a+b) à la puissance n revient à multiplier n binômes identiques (a+b). Le résultat est une somme où chaque élément est le produit de n facteurs de type a ou b choisi chacun dans un binôme différent. Les termes sont ainsi de la forme $a^{n-p}b^p$. Chacun de ces termes est obtenu autant de fois qu'il existe de façons de choisir les p éléments a parmi les n, c'est-à-dire le nombre de combinaisons C_n^p.

Compte tenu de la symétrie des combinaisons C_n^p la formule du binôme de Newton peut s'écrire :

$$\boldsymbol{(a+b)^n=\sum_{p=0}^{n} C_n^p a^{n-p} b^p=\sum_{q=0}^{n} C_n^q a^q b^{n-q}\ avec\ q=n-p}$$

Les **coefficients binomiaux**, C_n^p *ou* $\binom{n}{p}$ qui sont les coefficients de la formule du binôme de Newton figurent dans de nombreuses formules mathématiques, notamment pour le calcul des probabilités de la **loi binomiale**. Ces coefficients peuvent être obtenus facilement à l'aide du **triangle de Pascal.**

Exemple I.9 :

Le développement de $(a+b)^6$ donne :

$$(a+b)^6=\sum_{p=0}^{6}\binom{6}{p} a^{6-p} b^p$$

$$(a+b)^6=\binom{6}{0} a^6+\binom{6}{1} a^5 b+\binom{6}{2} a^4 b^2+\binom{6}{3} a^3 b^3+\binom{6}{4} a^2 b^4+\binom{6}{5} a b^5+\binom{6}{6} b^6$$

L'application du **triangle de Pascal** (7^{eme} ligne) donne directement les valeurs des coefficients binomiaux :

$$(a+b)^6=a^6+6a^5b+15a^4b^2+20a^3b^3+15a^2b^4+6ab^5+b^6$$

Remarque I.4 :

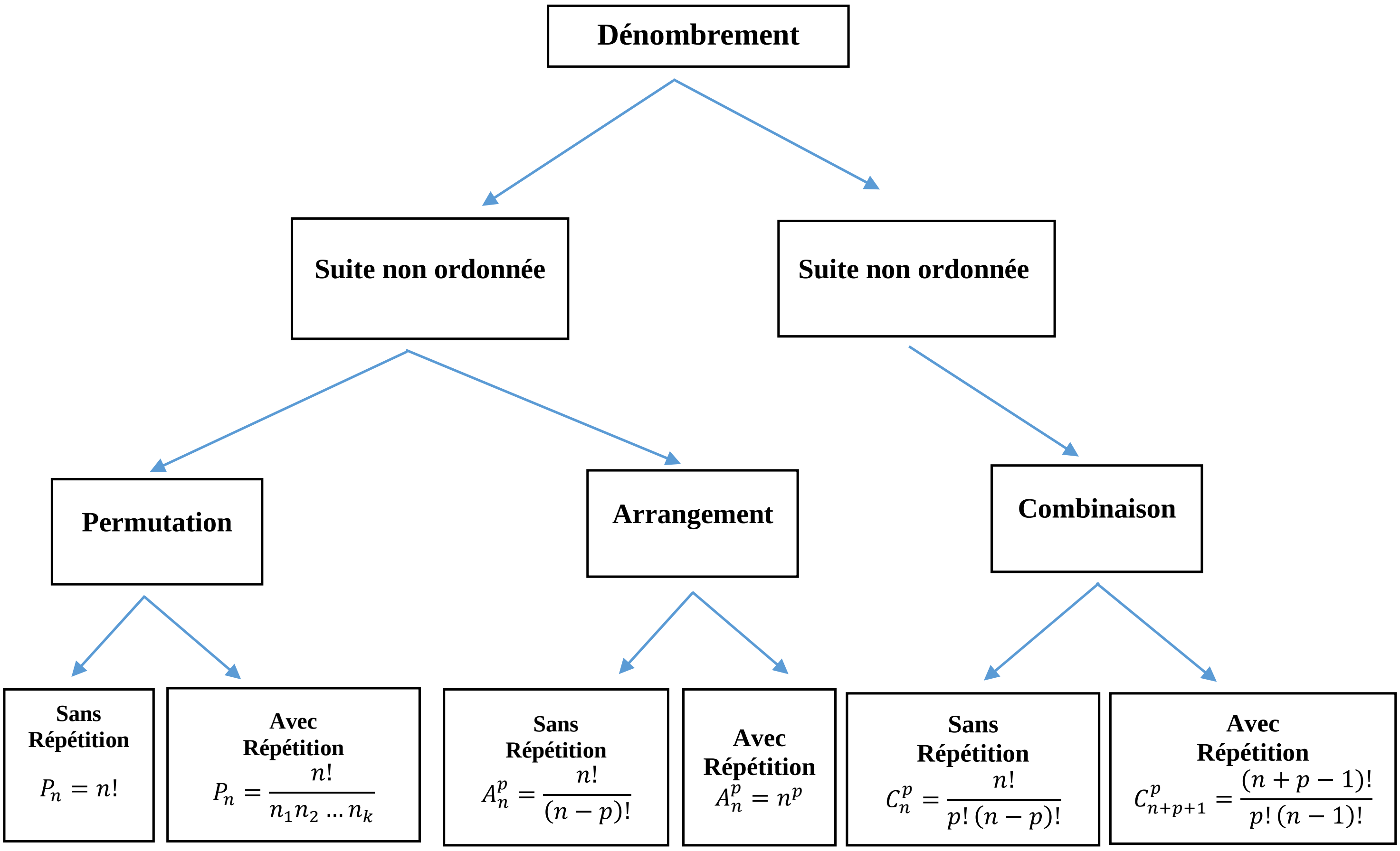
- Si l'on pose a = b = 1, on obtient alors, d'après la formule du binôme de Newton ;

$(2)^n = \sum_{p=0}^{n} C_n^P$; or C_n^P étant le nombre de parties à *p* éléments de l'ensemble *E* contenant *n* objets ; $\sum_{p=0}^{n} C_n^P$ représente le **nombre de parties ou partitions de l'ensemble *E*** que l'on note : $card\ E = n$; alors $card\ \mathcal{P}(E) = 2^n$

- Le **cardinal d'un ensemble** (card) correspond au nombre d'éléments constituant cet ensemble ;

RECAPITULATIF

Disposition	Ordre	Sans répétition	Avec répétition
Permutation	Oui	$P_n = n!$	$P_n = \dfrac{n!}{n_1 n_2 \dots n_k}$
Arrangement	Oui	$A_n^p = \dfrac{n!}{(n-p)!}$	$A_n^p = n^p$
Combinaison	Non	$C_n^p = \dfrac{n!}{p!\,(n-p)!}$	$C_{n+p+1}^p = \dfrac{(n+p-1)!}{p!\,(n-1)!}$



CHAPITRE II. CALCUL DES PROBABILITES

II.1. LE CONCEPTS D'EXPERIENCE ALEATOIRE ET D'EVENEMENT

II.1.1. Expérience aléatoire et ensemble fondamental des résultats

Le terme expérience aléatoire désigne une expérience dont on connaît l'ensemble des résultats possible sans savoir lequel de ceux-ci se réalisera. L'ensemble des résultats possible porte le nom **d'ensemble fondamental des résultats de l'expérience aléatoire**. On désigne souvent cet ensemble par la lettre grecque **Ω** (oméga majuscule).

II.1.2. Evènements

II.1.2.1. Définition

Tout sous-ensemble A de l'ensemble fondamental Ω des résultats d'une experience aléatoire porte le non d'**évènement**. On dit qu'**un évènement A** se réalise lorsque le résultat observé de l'expérience est un résultat appartenant au sous-ensemble A de Ω.

Exemple II.1 :

Considérons l'expérience aléatoire qui consiste à lancer un dé.

L'ensemble fondamental Ω de cette expérience aléatoire peut être donné par :
$\Omega = \{1, 2, 3, 4, 5, 6\}$.

L'évènement A défini par « le résultat du lancer est un nombre pair » correspond au sous-ensemble suivant $A = \{2, 4, 6\}$. Pour que cet évènement se réalise, il faut que le résultat du lancer soit un des trois nombres 2, 4, 6.

Exemple II.2 :

Si nous considérons l'expérience aléatoire qui consiste à lancer successivement deux dés, l'ensemble fondamental est formé de tous les couples de résultats possibles pour les deux dés ; nous dénombrons par conséquent 36 éléments :

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

En fonction de cet ensemble fondamental, nous pouvons par exemple décrire les événements suivants :

- **A** : La somme des points est égale à six :

$\mathbf{A} = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$;

- **B** : La somme des points est paire :

$\mathbf{B} = \{(1,1), (1,3), (1,5), (2,2), (2,4), (2,6), \dots (6,2), (6,4), (6,6)\}$;

- **C** : La somme des points est inférieure à six :

$\mathbf{C} = \{(1,1),(1,2),(3,3),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,2),(4,1)\}$.

II.1.2.2. Classification des évènements

- Un évènement A est dit **simple** s'il ne contient qu'un seul des résultats de Ω, c'est-à-dire l'événement ne peut pas être divisé en d'autres événements. Exemple $\{(1,3)\}$
- Un évènement A est dit **composé** s'il contient plus d'un des résultats de Ω;
- L'ensemble vide est bien un sous ensemble de Ω. On le dit événement **impossible** ;
- Ω est dit l'événement **certain**, c'est-à-dire Ω est toujours réalisé quel que soit l'issue de l'épreuve.
- Un événement A est dit **impossible** noté $\emptyset$ si il ne peut pas être réalisé quel que soit l'issue de l'épreuve
- Un événement est dit **complémentaire** ou **contraire** de l'évènement A, noté A^c *ou* $\bar{A}$ est l'événement qui est réalisé si et seulement si l'événement A ne l'est pas.

Exemple II.3 :

Reprenons l'exemple II.1, l'événement A « ensemble de nombres pairs » et A={2, 4, 6} son complémentaire est $A^c = \{1,3,5\}$

II.1.2.3. Opérations sur les évènements

Soit Ω l'ensemble fondamental. Si on considéré la réalisation de deux événements A et B, il est possible d'effectuer des opérations sur ces ensembles.

L'application des opérations intersection, réunion ou complémentation sur des événements définis à partir d'un même ensemble fondamentale Ω engendre d'autres événements. La terminologie utilisée en calcul des probabilités diffère légèrement de celle utilisée en théorie des ensembles. Le tableau suivant résume toutes les opérations sur les ensembles.

Evénement	Terminologie Ensembliste	Terminologie Probabiliste	Signification de la réalisation de l'événement étudié
$\bar{A}$	Complémentation de A	Négation de A	$\bar{A}$ se réalise si et seulement si A ne se réalise pas
$A \cap B$	Intersection de A et B	Conjonction de A et B	$A \cap B$ se réalise si et seulement si A **et** B se réalisent
$A \cup B$	Réunion de et B	Réunion de A et B	$A \cup B$ se réalise si et seulement si A **ou** B se réalisent

Exemple II.4 (intersection) :

On lance un dé, l'ensemble fondamental est $\Omega = \{1, 2, 3, 4, 5, 6\}$;

A « Obtenir un chiffre impair », $A = \{1, 3, 5\}$

B « Obtenir un chiffre inférieur à six », $B = \{1, 2, 3, 4, 5\}$;

$A \cap B = \{1, 3, 5\}$.

Exemple II.5 (union) :

On lance un dé, l'ensemble fondamental est $\Omega = \{1, 2, 3, 4, 5, 6\}$;

A « Obtenir un chiffre plus petit que 3 », $\mathrm{A} = \{1, 2\}$;

B « Obtenir un multiple de 3 », $\mathrm{B} = \{3, 6\}$;

$A \cup B = \{1, 2, 3, 6\}$.

II.1.2.4. Compatibilité de deux événements

Deux événements A et B sont **compatible** si les deux peuvent se réaliser en même temps. En d'autres mots, deux événements sont compatibles si leur conjonction est un événement possible, c'est-à-dire si $A \cap B \neq \emptyset$.

Deux événements A et B sont **incompatible** si $A \cap B = \emptyset$.

II.1.2.5. Quelques équivalences

La réalisation d'un événement A entraîne la réalisation d'un événement B si et seulement si $\subseteq B \; et \; A = \emptyset$, c'est-à-dire si et seulement si A est un événement possible inclus dans B.

La réalisation d'un événement A entraîne la non réalisation d'un événement B si et seulement si $A \cap B = \emptyset$, c'est-à-dire si et seulement si les événements A et B sont Incompatibles.

Soit A un événement possible ($A = \emptyset$), sans être certain ($A \neq \emptyset$). L'événement Ane se réalise pas si et seulement si sa négation (ou son complétement) $\bar{A}$ se réalise.

II.2. LE CONCEPT DE PROBABILITES

II.2.1. Probabilités de réalisation d'un événement lorsque Ω contient n résultats

Considérons une expérience aléatoire dont l'ensemble fondamental des résultats contient un nombre fini n de résultats, c'est-à-dire $\boldsymbol{\Omega} = \{\boldsymbol{\omega_1}, \boldsymbol{\omega_2}, \dots \boldsymbol{\omega_n}\}$.

Aucun des événements simples (évènements ne correspondant qu'à un seul résultat) qu'on peut définir pour cette expérience aléatoire correspond à une probabilité de réalisation ;

c'est-à-dire à chacun des événements $\{\boldsymbol{\omega_i}\}$ correspond une probabilité de réalisation $P(\{\omega_i\}), (i = 1, 2, 3, \dots n)$. Ces n probabilités sont telles que pour $i = 1, 2, 3, \dots, n$,

$$0 \leq P(\{\omega_i\}) \leq 1 \quad et \quad \sum_{i=1}^{n} P(\{\omega_i\}) = 1$$

Soit A un événement défini à partir de cette expérience, c'est-à-dire soit A un sous-ensemble de Ω. La probabilité de réalisation de l'événement A peut se noter P(A) et de façon générale, $P(A) = \sum_{\omega_i \in A} P(\{\omega_i\})$

II.2.2. Simplification du calcul lorsqu'il y a équiprobabilité

Si les n éléments de l'ensemble Ω sont **équiprobables**, c'est-à-dire :

Si $P(\{\omega_i\}) = 1/n \, , i = 1, 2, \dots$, n alors : $P(A) = \frac{\#A}{\#\Omega} = \frac{\#A}{n}$, Où le symbole # désigne la cardinalité (le nombre d'éléments) de l'ensemble qui le suit.

Exemple II.6 :

Exemple d'une situation où les éléments de Ω ne sont pas équiprobables.

Une urne contient 12 boules de forme et de masse identiques. Une d celles-ci est numérotée 2, six sont numérotées 3, trois sont numérotées 4 et les deux dernières sont numérotées 5. Considérons l'expérience aléatoire qui consiste à prélever au hasard une boule de l'urne et à noter son numéro. Définissons l'événement A par « le numéro de la boule prélevée est un nombre pair ».

Soit $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\} = \{2, 3, 4, 5\} et\ A = \{2, 4\}$.

Pour établir les probabilités de réalisation des 4 événements simples associables à Ω, on peut supposer que chacune des 12 boules de l'urne a la même probabilité d'être prélevée. On obtient alors que :

$P(\{\omega_1\} = P(\{2\}) = 1/12,$ $\qquad P(\{\omega_2\} = P(\{3\}) = 6/12$

$P(\{\omega_3\} = P(\{4\}) = 3/12\ ,$ $\qquad P(\{\omega_4\} = P(\{5\}) = 2/12$

Pour calculer la probabilité de la réalisation de l'événement A, on doit utiliser la formule générale de l'équation II.1, car il n'y a pas équiprobabilité de éléments de Ω. On obtient :

$P(A) = P(\{2,4\}) = P(\{2\}) + P(\{4\}) = 1/12 + 3/12 = 4/12.$

Exemple II.7 :

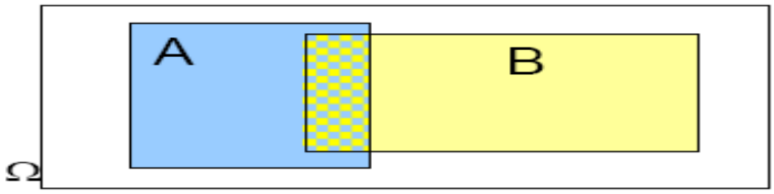
Exemple d'une situation où les éléments de Ω ne sont pas équiprobables.

Considérons la même expérience aléatoire que dans l'exemple II.2 pour une urne qui contient trois boules numérotées 2, trois boules numérotées 3, trois boules numérotées 4 et trois boules numérotées 5.

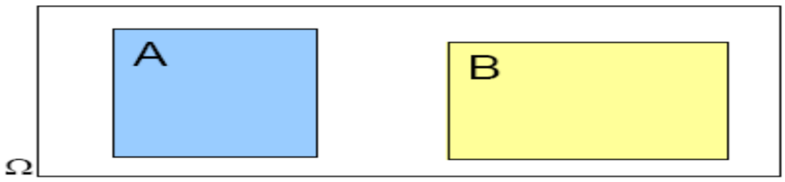
On est alors en présence d'un ensemble fondamental $\Omega = \{2, 3, 4, 4\}$ dont les 4 éléments sont équiprobables. Le calcul de la probabilité de la réalisation de l'événement $A = \{2, 4\}$ s'en trouve simplifié. Il suffit d'appliquer l'équation II.2 : $P(A) = \frac{\#A}{\#\Omega} = \frac{2}{4}$.

II.2.3. Quelques propriétés des probabilités

- $\mathrm{P}(\Omega) = 1$ et $\mathrm{P}(\emptyset) = 0$.
- Soit A et B deux événements définis à partir d'une même expérience aléatoire.
- $0 \leq P(A) \leq 1 \; et \; 0 \leq P(B) \leq 1$.
- $P(\bar{A}) = 1 - P(A) \; et \; P(\bar{B}) = 1 - P(B)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. (Voir l'illustration graphique suivante).



- Si A et B sont des événements incompatibles, c'est-à-dire si $A \cap B = \emptyset$, alors $P\,(A \cup B) = P(A) + P(B)$ (Voir l'illustration graphique suivante).



CHAPITRE III. CONDITIONNEMENT ET INDEPENDANCE

III.1. LE CONCEPT DE PROBABILITE CONDITIONNELLE

III.1.1. Définition

Soit A et E deux événements où E a une probabilité de réalisation non-nulle.

La Probabilité conditionnelle de réalisation de l'événement A étant donné que l'événement E s'est réalisé est notée $P(A|E)$.

Parfois, la probabilité conditionnelle $P(A|E)$ se calcule ou se déduit directement.

Lorsque ce n'est pas le cas, on utilise la définition : $P(A|E) = P$.

Exemple III.8 :

Soit l'expérience de jet d'un dé deux fois de suite, on définit A par « la somme des chiffres obtenus est inférieure ou égale à 4 ».

Calculer la probabilité d'avoir le chiffre 2 au $2^{\text{emé}}$ jet sachant que A est réalisé.

A « La somme des deux chiffres ≤ 4 ».

$$A = \{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\}, card(A) = 6\ ;$$

B « Avoir le chiffre 2 au $2^{\text{éme}}$ jet » ;

$$B = \{(1,2),(2,2),(3,2),(4,2),(5,2),(6,2)\}, card(B) = 6\ ;$$

$\Omega = \{(i,j), 1 \leq i,j \leq 6\}, card(\Omega) = 6^2$;

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

$P(A) = \frac{6}{36}$,

$A \cap B = \{(1,2),(2,2)\}, P(A \cap B) = \frac{2}{36}$;

$P(B|A) = \frac{P(A\cap b)}{P(A)} = \frac{\frac{2}{36}}{\frac{6}{36}} = \frac{1}{3}$.

Remarque III.3 :

La définition qui précédé ne doit pas donner la fausse impression que la probabilité de la conjonction (ou intersection) des deux événements est d'abord connue et qu'ensuite on peut calculer la probabilité conditionnelle. En pratique, c'est souvent l'inverse qui se produit : on connait d'abord les valeurs de $P(A|E)$ et $P\ (E)$, ce qui permet ensuite de calculer la probabilité de la conjonction (intersection) des deux événements : $P\ (A \cap E) = P(A|E)P(E)$.

III.1.2. Quelques propriétés des probabilités conditionnelles

Les propriétés des probabilités de réalisation de deux événements A et B s'étendent à la probabilité conditionnelle.

Soit A, B et E trois événements, où E a une probabilité de réalisation non nulle.

- $0 \leq P(A|E) \leq 1 \; et \; 0 \leq (B|E) \leq 1$;
- $P(\bar{A}|E) = 1 - P(A|E) \; et \; P(\bar{B}|E) = 1 - P(B|E)$;
- $P(A \cup B|E) = P(A|E) + P(B|E) - P(A \cap B|E)$;

Soit A et B sont des événements incompatibles, c'est-à-dire si $A \cap B = \emptyset$, alors : $P(A \cup B|E) = P(A|E) + P(B|E)$.

III.2. LE CONCEPT D'INDEPENDANCE EN PROBABILITE

III.2.1. Indépendance de deux événements

Deux événements possibles A et B sont dits **indépendants** en probabilité si la réalisation de l'un ne modifie en rien la probabilité de réalisation de l'autre. On peut définir ce concept d'indépendance en probabilité de deux événements possibles de trois façons équivalentes : $P(A|E) = P(A)$, $P(B|E) = P(B)$ ou $P(\cap B) = P(A)P(B)$.

III.2.2. Indépendance mutuelle de m événements

m événements A_1, A_2, A_3, …, A_m sont dits **mutuellement indépendants** en probabilité, si et seulement si la probabilité de réalisation de la conjonction (intersection) d'un nombre quelconque d'entre eux est toujours égal au produit des probabilités de réalisation de chacun.

Par exemple : trois événements A_1, A_2, A_3 dits **mutuellement indépendants** en probabilité, si et seulement si :

$P(A_1 \cap A_2) = P(A_1)P(A_2)$, $P(A_1 \cap A_3) = P(A_1)P(A_3)$, $P(A_2 \cap A_3) = P(A_2)P(A_3)$ et $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$.

Remarques III.4 :

Pour $m > 2$, il est possible de construire des exemples de m événements deux à deux indépendants en probabilité sans qu'ils soient indépendants dans leur ensemble.

En pratique, on peut parfois déduire directement l'indépendance mutuelle de me événements à partir de la façon dont est conçue l'expérience aléatoire. Dans un tel cas, la propriété d'indépendance mutuelle permet de simplifier le calcul de la probabilité de réalisation d'une conjonction de ces événements. En prend par exemple le tirage avec remise.

III.3. THEOREMES GENERAUX DE PROBABILITES

III.3.1. Théorème de multiplication

Théorème III.1 : $\forall\, A_1, A_2$ deux quelconque de Ω de probabilité non nulle.

$P(A_1 \cap A_2) = P(A_1) \times P(A_2/A_1) = P(A_2) \times P(A_1/A_2)$.

Démonstration : D'après la définition de probabilité conditionnelle on a :

$P(A_2/A_1) = \frac{P(A_1 \cap A_2)}{P(A_2)}$, $P(A_2 \cap A_1) = P(A_1) \times P(A_2/A_1)$.

Remarque III.5 :

$P(A_1/A_2) = P(A_1)$.

III.3.1.1. Généralité du théorème de multiplication

Théorème III.2 :

$\forall\, A_1, A_2, \dots, A_n$ deux quelconque de Ω de probabilité non nulle.

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

Démonstration :

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

$= \frac{PA_1}{1}\frac{P(A_1 \cap A_2)}{P(A_1)}\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \dots \frac{P(A_n \cap A_1 \cap A_2 \cap \dots A_{n-1}}{P(A_1 \cap A_2 \cap A_2 \cap \dots A_n} = P(A_1 \cap A_2 \cap \dots \cap A_{n-1}A_n$.

III.3.2. Théorème des probabilités totales

Théorème III.3 :

Soit une partition de Ω en événement $\{A_1, A_2, \dots, A_n\}$ de probabilités strictement positive, $P(A_i) > 0$ pour $1 \leq i \leq n$, et incompatibles deux à deux, i.e. avec $A_i \cap A_j = \emptyset$ pour $i \neq j$, et B un événement quelconque en Ω alors,

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

Exemple II. 9 :

Une boite N°1 contient 3 boules rouges et 2 blanches.

Une boite N°2 contient 2 boules rouges et 8 blanches.

On lance une pièce de monnaie régulière, si pile sort, une boule est alors choisie dans la 1ére boite, sinon elle est choisie dans la 2éme boite.

Trouver la probabilité qu'une boule rouge soit choisie.

Solution :

Soit B « Une boule rouge est choisie »

Soit A_i « La boule choisie provient de la boite i, i=1,2 »,

$P(A_1) = P(pile) = 1/2,$

$P(A_2) = P(face) = 1/2,$

$P(B/A_1) = \frac{3}{5}$, et $P(B/A_2) = 2/10,$

$P(B) = \frac{1}{2}\frac{3}{5} + \frac{1}{2}\frac{2}{10} = \frac{2}{5}.$

III.3.3. Théorème de Bayes

Théorème III.4 :

Soit (Ω, A, P) muni d'une partition $(A_1, A_2, \dots, A_n)$, alors pour tout événement B et pour tout indice i on a :

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

Démonstration :

Il suffit en effet d'écrire :

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

Puis d'utiliser la décomposition $P(B \cap A_i) = P = P(A_i)P(B/A_i)$ pour le numérateur et la formule des probabilités totale pour le dénominateur.

Exemple II. 10 :

Soit trois urnes telles que :

- Urne N°1 : Contient 3 boules rouges et 5 noires ;
- Urne N°2 : Contient 2 boules rouges et 1 noire ;
- Urne N°3 : Contient 2 boules rouges et 3 noires ;

On prend une urne au hasard et on tire boule. Si la boule tirée est rouge, quelle est la probabilité pour qu'elle provienne de l'urne N°1 ?

Solution :

B « La boule tirée est rouge »,

A_i « La boule tirée provient de l'urne i, i=1, 2, 3 ».

$P(A_1) = P(A_2) = P(A_2) = \frac{1}{3},$

$P(B/A_1) = \frac{3}{8}, \quad P(B/A_2) = \frac{2}{3}, \quad P(B/A_3) = \frac{2}{5}.$

$P(A_1/B) = \frac{P(A_1)P(B/A_1)}{\sum_{i=1}^{3} P(A_i)P(B/A_i)} = \frac{\frac{13}{38}}{\frac{13}{38}+\frac{12}{33}+\frac{12}{35}} = 0.26\ .$

CHAPITRE IV. VARIABLES ALEATOIRES

En 1654, Blaise Pascal (1623 ; 1662) entretient avec Pierre de Fermat (1601 ; 1665) des correspondances sur le thème des jeux de hasard et d'espérance de gain qui les mènent à exposer une théorie nouvelle : les calculs de probabilités.
Ils s'intéressent à la résolution de problèmes de dénombrement comme par exemple celui du Chevalier de Méré : «Comment distribuer équitablement la mise à un jeu de hasard interrompu avant la fin ?»

IV.1. DEFINITIONS ET PROPRIETES

Dans la plupart des phénomènes aléatoires, le résultat d'une épreuve peut se traduire par une « grandeur » mathématique, très souvent représentée par un nombre entier ou un nombre réel. La notion mathématique qui représente efficacement ce genre de situation concrète est celle de variable aléatoire (notée également v.a.). Ainsi le temps de désintégration d'un atome radioactif, le pourcentage de réponses « oui » à une question posée dans un sondage ou le nombre d'enfants d'un couple sont des exemples de variables aléatoires.

Définition IV.1 :

Une **variable aléatoire X** est une fonction de l'ensemble fondamental Ω à valeurs dans R, $X : \Omega \to R$.

Lorsque la variable X ne prend que des valeurs discrètes, on parle **de variable aléatoire discrète**.

Un **vecteur aléatoire** $X : \Omega \to R^d$ est une fonction $X = (X_1, \ldots, X_d)$ à valeurs dans R^d telle que les coordonnées X_i soient des variables aléatoires.

Pour tout intervalle $[a, b] \subset R$, l'ensemble $\{X \in [a, b]\} = \{\omega \in \Omega : X(\omega) \in [a, b]\}$ est **un événement**.

Exemple IV.1 :

1. On jette deux dés distincts et on s'intéresse à la somme des points. On note X cette variable aléatoire, elle est définie par :

$$X : \Omega \to R \text{ avec } \Omega = \{(1, 1), (1, 2), \ldots, (6, 5), (6, 6)\}$$
$$(\omega_1, \omega_2) \to \omega_1 + \omega_2.$$

L'ensemble des valeurs possibles de X est $\{2, 3, \ldots, 12\}$.

2. On lance toujours deux dés, mais cette fois on s'intéresse au plus grand chiffre Y obtenu. On a alors

$$Y : \Omega \rightarrow R \text{ avec } \Omega = \{(1, 1), (1, 2), \ldots, (6, 5), (6, 6)\}$$
$$(\omega_1, \omega_2) \rightarrow \max(\omega_1, \omega_2).$$

La variable Y est à valeurs dans $\{1, 2, \ldots, 6\}$.

3. On observe deux bactéries et on s'intéresse à la durée de vie T de la bactérie qui disparaîtra la première. L'ensemble fondamental est $\Omega = [0, +\infty[\times[0, +\infty[$. La variable T s'écrit alors :

$$T : \Omega \longrightarrow R$$
$$(\omega_1, \omega_2) \rightarrow \inf\{\omega_1, \omega_2\}.$$

Seul le dernier exemple n'est pas une variable discrète.

IV.2. LOI DE PROBABILITE, FONCTION DE REPARTITION

La loi de probabilité d'une variable aléatoire permet de connaitre les chances d'apparition des différentes valeurs de cette variable.

On se place sur l'espace de probabilité $(\Omega, \mathbb{P})$.

Définition IV.2 :

Soit X une variable aléatoire. La loi de probabilité de X est définie par la fonction FX, appelée fonction de répartition de la variable X, définie par :

$$F_X : R \rightarrow [0, 1]$$
$$x \rightarrow \mathbb{P}(X \leq x).$$

On dit que deux variables aléatoires X et Y ont la même loi si elles ont la même fonction de répartition $F_X = F_Y$.

Remarque IV.1 :

Soit I un intervalle de R. L'événement $\{X \leq x\}$ représente l'ensemble des valeurs $\omega \in \Omega$ telles que $X(\omega)$ soit inférieur à x, i.e. $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}$.

La loi de X est en générale notée L(X) ou Loi(X).

Remarque IV.2 :

On a $\mathbb{P}(X \in R) = 1$, car $\mathbb{P}(X \in R) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in R\}) = \mathbb{P}(\Omega) = 1$.

Propriétés IV.1 :

La fonction de répartition est une fonction croissante à valeur dans [0, 1] telle que $\lim_{x\to-\infty} F_X(x) = 0 \quad et \quad \lim_{x\to+\infty} F_X(x) = 1$ mais elle n'est pas forcément continue.

Remarque IV.3 :

Soit $a \leq b$, on a $\mathbb{P}(X \in [a, b]) = \mathbb{P}(X \leq b) - \mathbb{P}(X < a)$.

IV.2.1. Loi d'une variable discrète

La fonction de répartition d'une variable discrète est constante par morceaux. Si X est une variable discrète à valeurs dans $\{x_1, \ldots, x_n\}$ avec $x_1 < \ldots < x_n$ alors pour $x \in R$:

$$F_X(x) = \sum_{I=1}^{K} \mathbb{P}(X = x_i) \text{ avec k tel que } x_k \leq x < x_{k+1}.$$

De même, si X prend une infinité de valeurs $\{x_1, \ldots, x_n, \ldots\}$ avec $x_1 < \ldots < x_n \ldots$, on a pour $x \in R$

$$F_X(x) = \sum_{I=1}^{K} \mathbb{P}(X = x_i) \text{ avec k tel que } x_k \leq x < x_{k+1}.$$

Les sauts de la fonction de répartition F_X ont lieu en les points xi et la hauteur du saut au point xi est égale à $\mathbb{P}(X = x_i)$. Il suffit donc de calculer la fonction de répartition aux points x_i.

Proposition IV.1

Si X est à valeurs discrètes dans $\{x_1, \ldots, x_n\}$ (ou $\{x_1, \ldots, x_n, \ldots\}$), la loi de X est entièrement caractérisée par $\{\mathbb{P}(X = x_i) : i \geq 1\}$.

On remarque que

1. pour tout $i \geq 1$, $\mathbb{P}(X = x_i) \in [0, 1]$,
2. $\sum_{i \geq 1} \mathbb{P}(X = xi) = 1$. (En effet, $1 = \mathbb{P}(X \in R) = \sum_{i \geq 1} \mathbb{P}(X = x_i)$).

Exemple IV.2 :

Reprenons les deux premiers exemples précédents, qui décrivent effectivement des variables discrètes. Pour trouver la loi de ces variables on utilise la proposition IV.1.

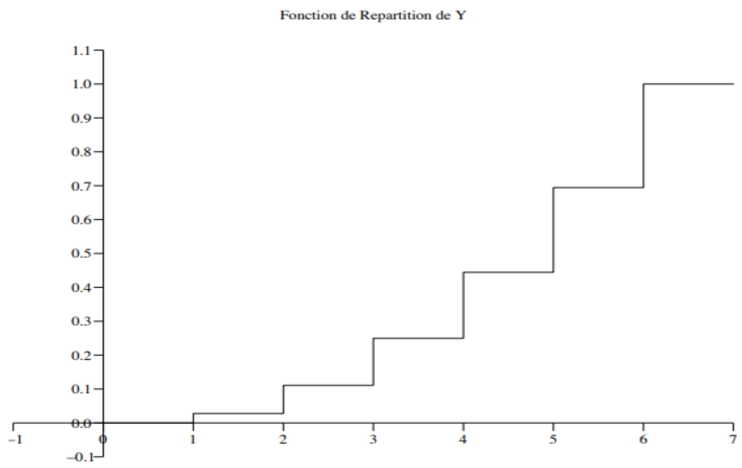
1. X est à valeurs dans $\{2, 3, \ldots, 12\}$, donc $\mathbb{P}(X = k) = 0$ pour $k \notin \{2, 3, \ldots, 12\}$.

$\boldsymbol{k}$	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(\boldsymbol{X = k})$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2. Y est à valeurs dans $\{1, 2, \ldots, 6\}$, donc $\mathbb{P}(Y = k) = 0$ pour $k \notin \{1, 2, \ldots, 6\}$.

$\boldsymbol{k}$	1	2	3	4	5	6
$\mathbb{P}(\boldsymbol{Y = k})$	1/36	3/36	5/36	7/36	9/36	11/36
$\boldsymbol{F_Y(k)}$	1/36	4/36	9/36	16/36	25/36	1

La fonction de répartition de Y est :



IV.3. ESPERANCE MATHEMATIQUE

L'idée intuitive de l'espérance puise son origine dans les jeux de hasard. Considérons le jeu suivant : on lance un dé plusieurs fois de suite. Supposons que pour une mise de 1 euro, on gagne 1 euro si le résultat obtenu est pair, 2 euros si le résultat est 1 ou 3, et on perd 3 euros si le résultat est 5. Est-il intéressant de jouer à ce jeu ? Quel peut-être le gain moyen ?

Soit X la variable aléatoire correspondant au nombre d'euros gagnés ou perdus. La loi de X est :

k	-3	1	2
$\mathbb{P}(X=k)$	1/6	1/2	1/3

L'espérance de gain, noté $\mathbb{E}$ [X], est alors $\mathbb{E}$ [X] = −3, 1/6+1, 1/2+2, 1/3 = 2/3. Le joueur gagne donc en moyenne 2/3 d'euros pour une mise de 1 euro. . .

Définition IV.3 :

L'espérance d'une variable aléatoire X est notée $\mathbb{E}[X]$. Elle représente la valeur moyenne prise par la variable X.

1. Si X est une variable discrète à valeurs dans $\mathbb{D} = \{x_1, \ldots, x_n\}$, son espérance est : $\mathbb{E}[X] = x_1\mathbb{P}(X = x_1)+ \ldots + x_n\mathbb{P}(X = x_n) = \sum_{i=1}^{n} x_i\mathbb{P}(X = x_i)$
2. Si X est une variable discrète à valeurs dans l'ensemble infini $\mathbb{D} = \{x_i : i \geq 1\}$, lorsque la somme est bien définie, son espérance est : $\mathbb{E}[X] = \sum_{i=1}^{+\infty} x_i\mathbb{P}(X = x_i)$
3. Si X est une variable à densité f, lorsque l'intégrale est bien définie, son espérance est : $\mathbb{E}[X] = \int_{-\infty}^{+\infty} x\, f(x)dx$.

Lorsqu'une variable X vérifie $\mathbb{E}[X] = 0$, on dit que la variable est centrée.

Propriétés IV.2.

1. L'espérance est linéaire : Soient a et b ∈ R, deux variables aléatoires X et Y d'espérance finie alors : $\mathbb{E}[aX + bY] = a\,\mathbb{E}[X] + b\,\mathbb{E}[Y]$;
2. Si $X \geq 0$, alors $\mathbb{E}[X] \geq 0$;
3. Si $X \leq Y$, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Exemple IV.3 :

Supposons que la durée de vie T d'une bactérie est modélisée par la loi exponentielle de densité $f(t) = \lambda \exp(-\lambda t)$ pour $t \geq 0$ pour une certaine valeur de λ. Alors sa durée de vie moyenne est $\mathbb{E}(T) = 1/\lambda$.

IV.3.1. Espérance d'une fonction

On considère une variable aléatoire X et une fonction h. On aimerait connaitre l'espérance de la nouvelle variable $Y = h(X)$.

Proposition IV.2 :

Soit X une variable aléatoire, h une fonction définie sur R. Alors :

1. Si X est une variable discrète à valeurs dans $\mathbb{D} = \{x_1, \ldots, x_n\}$, on a

$\mathbb{E}[h(X)] = h(x_1)\mathbb{P}(X = x_1) + \ldots + h(x_n)\mathbb{P}(X = x_n) = \sum_{i=1}^{+\infty} h(x_i)\mathbb{P}(X = x_i)$;

2. Si X est une variable discrète à valeurs dans l'ensemble infini $D = \{x_i : i \geq 1\}$, lorsque la somme est bien définie, on a : $\mathbb{E}[h(X)] = \sum_{i=1}^{+\infty} h(x_i)\mathbb{P}(X = x_i)$;
3. Si X est une variable à densité f, lorsque l'intégrale est bien définie, on a

$$\mathbb{E}[h(X)] = \int_0^{+\infty} h(x)f(x)dx$$

IV.4. MESURER LA DIFFUSION D'UNE VARIABLE : VARIANCE ET ECART TYPE

On a vu que l'espérance correspondait à la valeur moyenne d'une variable aléatoire. L'écart type représente l'écart moyen (la distance moyenne) entre la variable et sa moyenne. Elle mesure la dispersion d'une variable, plus l'écart-type est grand plus la variable prend des valeurs qui peuvent être éloignées les unes des autres, plus l'écart-type est petit plus la variable prend des valeurs proches de sa moyenne.

Définition IV.4 :

La variance d'une variable aléatoire X, notée Var(X), est définie par :

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$$

L'écart type est la racine carrée de la variance :

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

Lorsqu'une variable X vérifie Var(X) = 1, on dit que la variable est réduite.

Remarque IV.4 :

La variance s'écrit aussi $Var(X) = \mathbb{E}\,[X^2] - \mathbb{E}[X]^2$.

Propriétés IV.3 :

1. Var(X) = 0 si X est constante.
2. Soient a et b ∈ R, alors $Var(aX + b) = a^2 Var(X)$.

Exemple IV.4 :

Supposons que la durée de vie T d'une bactérie est modélisée par la loi exponentielle de densité $f(t) = \lambda \exp(-\lambda t)$ pour $t \geq 0$ pour une certaine valeur de λ. La variance de la durée de vie de la bactérie étudiée est $Var(T) = 1/\lambda^2$.

CHAPITRE V. LOIS DE PROBABILITE DISCRETES USUELLES

V.1. LOI ET VARIABLE DE BERNOULLI

V.1.1. Définition

Soit une épreuve aléatoire comportant deux issues, deux événements élémentaires appelés souvent succès et échec dont les probabilités respectives sont p et q avec p + q = 1.

On définit alors : Ω= {S, E}

Avec :

p(S) = p et p (E) = q. Soit la variable aléatoire X : Ω ⟶ {0, 1} telle que X(S) = 1 ; X(E) = 0. La variable X est appelée variable de Bernoulli dont la loi de probabilité est

x_i	0	1	Total
p_i	q	p	1
$p_i x_i$	0	p	E(X) = p
$p_i x_i^2$	0	p	$E(X^2)$ = p

On note cette loi

$$X \rightsquigarrow \mathcal{B}(p)$$

Remarque V.1.1 :

Cette loi ne dépend que d'un paramètre p, la probabilité de succès.

V.1.2. Moments

Espérance : $E(X) = p$

Variance :$V(X) = p - p^2 = p(1 - p) = pq$

Ecart-type : $\delta = \sqrt{pq}$

Exemple V.1 :

Un entreprise possède 10 chaînes de fabrication C_1, C^2,. . . ,C_{10}. Elle sait qu'une chaîne possède un problème mais elle ignore laquelle, elle choisit alors une chaîne au hasard. On considère la variable aléatoire X prenant la valeur 1 si la chaîne testée est la chaîne défaillante et 0 sinon. Dans ce cas X est une variable de Bernoulli de paramètre 1/10. L'espérance et la variance de cette variable valent respectivement E(X) = 0, 1 et V (X) = 0, 09.

V.2. LOI ET VARIABLE BINOMIALES

V.2.1. Définition

Soit une épreuve de Bernoulli. À n répétitions indépendantes de cette épreuve de Bernoulli sont associées n variables aléatoires X_1, X_2,. . . ,X_n indépendantes.

On considère la variable aléatoire $X = X_1 + X_2 + . . . + X_n$. Cette variable X désigne le nombre de succès lors de n épreuves. L'univers image de la variable X est $\{0, 1, 2, . . . , n\}$. On a :

$$P(\{X = k\}) = C_n^k p^k q^{n-k}$$

Preuve : L'événement $\{X = k\}$ est obtenu par le résultat de k succès et $n \# k$ échecs. On peut avoir par exemple : $\underbrace{S \dots S}_{k\, fois} \underbrace{E \dots E}_{n-k\, fois}$ de probabilité $p^k q^{n-k}$, mais il existe C_n^k événements comportant k succès et $(n \# k)$ échecs d'où le résultat.

La variable ainsi définie suit une **loi binomiale** et on note :

$$X \rightsquigarrow \mathcal{B}(n, p)$$

On a bien défini une loi de probabilité puisque pour $p \in]0; 1[, C_n^k p^k (1-p)^{n-k} \geq 0, \forall k \in \{0, 1, \dots, n\}$ et $\sum_{k=1}^{n} p^k (1-p)^{n-k} = p + (1-p) = 1$ d'après la formule du binôme.

Remarque V.2.1 :

Si une variable aléatoire X représente le nombre de succès dans une série de n expériences de Bernoulli identiques et indépendantes alors X suit une loi binomiale de paramètres n et p où p représente la probabilité de succès lors d'une épreuve de Bernoulli.

V.2.2. Moments

Espérance : $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i) = np$

Variance : $V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i) = npq$

Ecart-type : $\delta = \sqrt{npq}$

Remarque V.2.2 :

On a le rapport $\frac{p(\{X=k+1\})}{p(\{X=k\})} = \frac{C_n^{k+1} p^{k+1} q^{n-1-k}}{C_n^k p^k q^{n-k}} = \frac{n-k}{k+1} \times \frac{p}{q}$ d'où

$p(\{X = k + 1\}) = \frac{n-k}{k+1} \times \frac{p}{q} \times p(\{X = k\}$ relation qui permet de disposer de $p(\{X = k + 1\})$ lorsqu'on a déjà $p(\{X = k\})$.

Exemple V.2.1 :

Dans une population très nombreuse, on estime que la probabilité pour qu'une personne soit atteinte d'une maladie donnée est 0, 1. On choisit au hasard 1000 personnes de cette population (avec l'éventualité de choisir plusieurs fois la même personne). On note X la variable aléatoire représentant le nombre de personnes atteintes de la maladie parmi les 1000. X représente le nombre de succès (c'est-à-dire être atteint par la maladie) dans une suite de 1000 épreuves de Bernoulli (la personne est atteinte ou pas) identiques et indépendantes donc :

$$X \rightsquigarrow \mathcal{B}\,(1000;\ 0,1).$$

V.3. LOI ET VARIABLE DE POISSON

On rappelle la formule : $\lim\limits_{n\to+\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = \sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x$

1.3.1. Définition

On dit que la variable aléatoire X suit une loi de Poisson de paramètre $\lambda(\lambda \in \mathbb{R}^{+*}$ si et seulement si l'univers image de X est ℕ et

$$p(\{X = k\}) = \frac{e^{-\lambda}\lambda^k}{k!}$$

On note cette variable $\boldsymbol{X} \rightsquigarrow \mathcal{P}\,(\boldsymbol{\lambda})$

Puisque $\sum_{i\in\mathbb{N}} p_i = \sum_{i\in\mathbb{N}} e^{-\lambda}\frac{\lambda^i}{i!} = e^{-\lambda}\sum_{i\in\mathbb{N}}\frac{\lambda^i}{i!} = e^{-\lambda}e^{\lambda} = 1\ et\ \forall\ k \in \mathbb{N}, \frac{e^{-\lambda}\lambda^k}{k!} \geq 0,$ on a bien une loi de probabilité.

Remarque V.3.1

La loi de Poisson est encore appelée **loi des événements rares**. On peut admettre que le nombre de pannes survenant sur une machine donnée au cours d'une période donnée suit une loi de Poisson, ou encore le nombre d'accidents qui se produisent à un carrefour donné pendant une période donnée.

V.3.2. Les moments

Espérance : $E(X) = \lambda$

Variance : $V(X) = \lambda$

Ecart-type : $\delta = \sqrt{\lambda}$

Exemple V.3.1 :

Dans une entreprise. on admet que le nombre de pièces défectueuses produites par minute est une variable aléatoire X avec X $\rightsquigarrow \mathcal{P}\,(3)$. Déterminons la probabilité de l'événement A : "le nombre de pièces défectueuses produites en 1 minute est supérieur à 3". Par définition, $p(A) = p(\{X > 3\}) = 1 - p(\{X \leq 3\}) = 1 - 0{,}6472 = 0{,}3528$.

CHAPITRE VI. LOIS DE PROBABILITE CONTINUES USUELLES

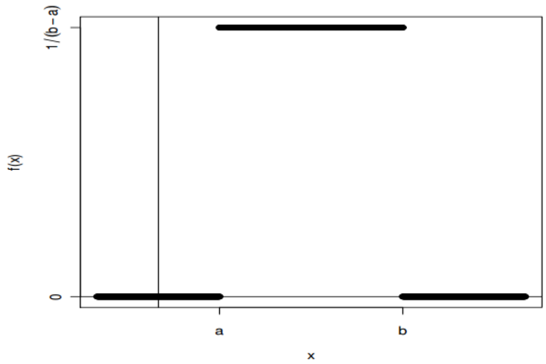
VI.1. LOI ET VARIABLE UNIFORMES

VI.1.1 Définition

On dit que la loi de probabilité d'une variable aléatoire réelle est **uniforme** sur un segment $[a; b]$, avec $0 \leq a < b$, si sa densité de probabilité f est définie par :

$$f(x) = \begin{cases} \dfrac{1}{b-a} & pour\ x \in [a; b] \\ 0 & pour\ < a\ ou\ x > b \end{cases}$$

On note alors $X \rightsquigarrow \mathcal{U}([a;\ b])$. f admet la représentation graphique suivant :



On a bien une densité de probabilité puisque :

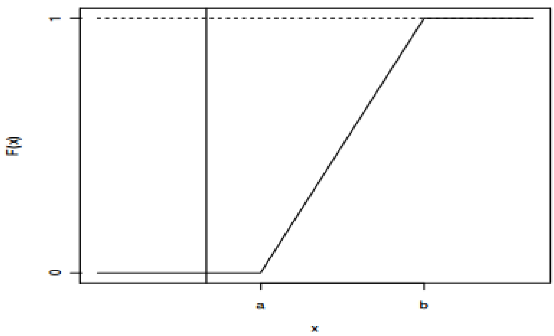
- $f(x) \geq 0\ \forall\ x \in \mathbb{R}$
- f est continue sur $]-\infty\ ;\ a[\cup]a\ ;\ b[\cup]b\ ;\ +\infty[$
- $\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{a} f(t)dt = \int_{a}^{b} f(t)dt = \int_{b}^{+\infty} f(t)dt = 0 + 1 + 0 = 1.$

VI.1.2 Fonction de répartition

1. On sait que $F(x) = p(\{X \leq x\}) = \int_{-\infty}^{x} f(t)dt\ \ donc\ si\ X \rightsquigarrow \mathcal{U}([a;\ b])$

$$F(x) = \begin{cases} 0 & pour\ x < a \\ \dfrac{x-a}{b-a} & pour\ a \leq x \leq b \\ 1 & pour\ x > b \end{cases}$$

La représentation de cette fonction est représenté sur le graphique suivant :



VI.1.3. Moments

Si $X \rightsquigarrow \mathcal{U}([a;\ b])$ alors

$$E\ (x) = \int_{-\infty}^{+\infty} tf(t)dt = \int_a^b \frac{t}{b-a}dt = \left[\frac{t^2}{2(b-a)}\right]_a^b = \frac{a+b}{2}$$

$$V(X) = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(a+b)^2}{12}$$

Exemple VI.1.1. On considère la fonction *f* définie par :

$$f(x) = \begin{cases} 0 & pour\ x < 0\ ou\ x > 1 \\ 1 & pour\ 0 \le X \le 1 \end{cases}$$

Il apparaît en intégrant que :

$$F(x) = \begin{cases} 0 & pour\ x < a \\ x & pour\ 0 \le x \le 1 \\ 1 & pour\ x > b \end{cases}$$

On trouve $E(X) = 1/2$ et $V(X) = 1/12$.

Remarque VI.1.1 :

Dans l'exemple précédent, comme dans toute variable aléatoire absolument continue, on a :

$$p(\{X = 0\}) = 0.en\ effet, \forall\ a \in \mathbb{R}, \int_a^b f(t)dt = 0.$$

VI.2. LOI DE LAPLACE-GAUSS OU LOI NORMALE

VI.2.1. Définition

On appelle **variable aléatoire normale** ou **gaussienne** toute variable aléatoire absolument continue dont la densité de probabilité *f* est définie par

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

m étant une constante réelle, " une constante réelle strictement positive. On utilise la notation suivante :

$$X \rightsquigarrow \mathcal{N}(m, \sigma).$$

Remarque VI.2.1 :

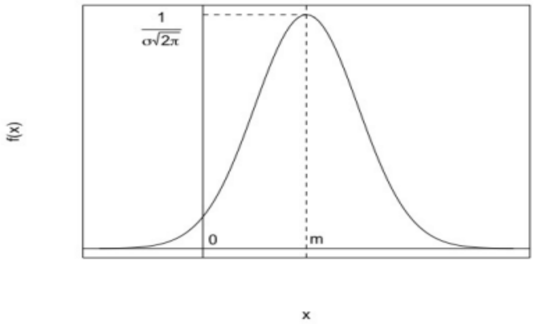
On admettra que f est bien une densité de probabilité (la difficulté étant de montrer que : $\int_{-\infty}^{+\infty} f(t)dt = 1$.

IV.2.2 Représentation graphique

La courbe représentative de *f* est donnée par le graphique suivant :

Remarque VI.2.2 :

- La courbe, dite courbe en cloche, a un axe de symétrie qui est la droite d'équation $x = m$.
- La densité *f* a un maximum atteint pour $x = m$ valant $\frac{1}{\sigma\sqrt{2\Pi}}$.
- La courbe est d'autant plus pointue que σ est petit.



VI.2.3. Moments

L'espérance et la variance d'une variable normale sont respectivement données par :

$$E(X) = m \text{ et } V(X) = \sigma^2$$

VI.3. EXPONENTIELLE

VI.3.1 Définition

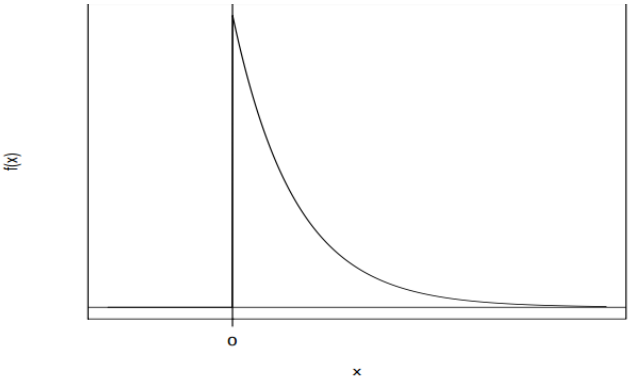
Une variable aléatoire X suit une loi exponentielle de paramètre $\lambda(\lambda \in \mathbb{R}^{+*})$ si X est une variable aléatoire absolument continue dont la densité de probabilité est définie par :

$$f(x) = \begin{cases} 0 & pour\ x < 0 \\ \lambda e^{-\lambda x} & pour\ x \geq 0 \end{cases}$$

On note alors

$$X \rightsquigarrow \varepsilon(\lambda).$$

La fonction f admet la représentation graphique suivante :



On a bien une densité de probabilité puisque :

$f(x) \geq 0 \; \forall \; x \in \mathbb{R}$,

f est continue sur $\mathbb{R}^{+} \; et \; \mathbb{R}^{-*}$,

$\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{0} 0dt + \int_{0}^{+\infty} \lambda e^{-\lambda t}\, dt. Or, si\ A > 0, \int_{0}^{A} \lambda e^{-\lambda t}\, dt = \left[-e^{-\lambda t}\right]_{0}^{A} = 1 - e^{-\lambda A} \;\rightarrow 1$

Quand $A \rightarrow \; +\infty$ donc $\int_{-\infty}^{+\infty} f(t)dt = 1$.

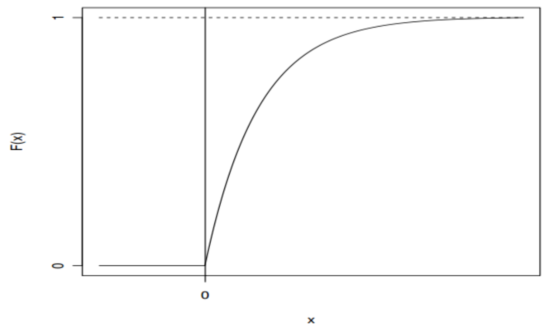
La loi exponentielle peut être considérée comme l'équivalent en continu de la loi géométrique dans le cas discret. En effet, elle modélise un temps d'attente du premier succès dans un processus de Poisson.

VI.3.2. Fonction de répartition

Si X $\rightsquigarrow$ $\varepsilon(\lambda)$, on a :

$$f(x) = \begin{cases} 0 & pour\ x < 0 \\ 1 - \lambda e^{-\lambda x} & pour\ x \geq 0 \end{cases}$$

Dont la représentation graphique est donnée par le graphique suivant :



VI.3.3. Les moments

Si X $\rightsquigarrow$ $\varepsilon(\lambda)$, on a :

- $E(X) = \int_{-\infty}^{+\infty} tf(t)dt = \int_{0}^{+\infty} \lambda t e^{-\lambda t}\, dt = \frac{1}{\lambda}$ à l'aide d'une intégration par parties.

- $V(X) = E(X)^2 - (E(X))^2$ Comme $E(X)^2 = \int_0^{+\infty} \lambda t^2 e^{-\lambda t}\, dt$ à l'aide de deux intégrations par parties, on obtient $E(X^2) = \frac{2}{\lambda^2}$