Chapitre 1

# INTRODUCTION A LA METHODE STATISTIQUE

## I-INTRODUCTION ET TERMINOLOGIE :

Il est nécessaire, avant d'exposer les différentes méthodes utilisées, de donner la terminologie c'est-à-dire la définition des termes statistiques qui vont êtres utilisés le long de ce cours. Nous devons tout d'abord faire la distinction entre les deux notions suivantes :

- STATISTIQUES : (avec s, au pluriel) désigne tout ensemble de données chiffrées relatives à un phénomène donné et recueillies en générale par des organismes spécialisés.

Par exemple : INS : Institut National de la Statistique centralise toutes les informations statistiques publiques et même privée. On y trouve toutes les données sociodémographiques, économiques et financières du pays. Il publie par exemple : l'économie de la Tunisie en chiffres ce sont les renseignements relatifs aux prix, salaires, production agricoles, échanges extérieurs, consommation….

On remarque bien que ces statistiques sont importantes et nécessaires pour le travail d'un statisticien mais elles sont insuffisantes d'où la définition de la deuxième notion à savoir :

- STATISTIQUE : (sans s, au singulier) est la science et l'ensemble des procédés avec lesquels on va pouvoir étudier les statistique**s ;** afin de répondre à certaines questions relatives à un ou plusieurs phénomène étudié. Donc la statistique est l'outil de travail de la matière première constituée par les statistique**s**.

On appellera données statistiques un ensemble de mesure observées sur une population donnée relative à un ou plusieurs caractères.

### Section 1 : DEFINITION ET TERMINOLOGIE

**1)** population :

Une population est un ensemble d'individus ou d'unités statistiques. Une population au sens statistique, n'est pas nécessairement un ensemble d'être humains. Elle peut être constituée de n'importe quel ensemble d'objets concernés par l'étude.

a) **Exemples :**

La population des ménages d'une cité, des vaches laitières en Tunisie, des étudiants d'une faculté.

- Une population peut être exhaustive c'est-à-dire couvrir l'ensemble de touts les individus concernés comme elle peut être une partie de cet ensemble, dans ce cas on parle d'échantillon. Si on revient à l'exemple des étudiants d'une faculté, les étudiants de la première année de cette faculté représentent un échantillon de la population totale.

**2)** unité statistique ou individu :

C'est un seul élément de l'ensemble de la population. Une population est donc composée de plusieurs unités statistiques ou individus.

b) **Remarque :**

Une population comporte toujours des unités homogènes (de même type, même nature) dont le nombre est fini. Une population ne peut pas comportée en même temps des voitures et des vaches.

**3)** Caractère – Modalité :

Un caractère est un aspect observable du phénomène étudié : c'est une dimension du phénomène. Une unité statistique peut être observée selon plusieurs caractères.

**Exemple :** P : population de voitures circulant à Nabeul ;
U : unité statistique : une voiture parmi ces voitures ;
C : caractères : âge ; marque ; maison ; prix ; puissance……

Chaque caractère se définit par l'ensemble des modalités qui sont les différentes valeurs possibles ou les différents états possibles ou les différentes situations possibles du caractère.

**Exemple :** P : population des étudiants de la $1^{ère}$ année de FESG Nabeul ;
U : Individu : un étudiant de la $1^{ère}$ année de FESG Nabeul ;
C : caractères : Sexe ; M : modalité : M ; F (deux)
Moyenne en BAC : M : modalité : tout l'intervalle [9,20] (infinité)
Age en année : M : modalité : 17-18-19-20-21-22-23-24-25 ( neuf)

On remarque que les modalités sont soient des états soient des chiffres, ce qui nous amène à dire qu'il y a deux types de caractères : qualitatifs et quantitatifs.

a) **Caractères qualitatifs :** ce sont les caractères dont les modalités diffèrent par leur nature. Ces modalités ne peuvent être mesurée, elle sont plutôt identifiées ou constatées. Comme par exemple les caractères : sexe ; couleur ; marque….. la liste des modalités est appeler Nomenclature.
b) **Caractères quantitatifs :** ils possèdent des modalités mesurables chiffrées telle que par exemple : taille, âge, revenu…….on distingue :
   - Caractères quantitatifs discrets : c'est un caractère qui ne peut prendre que des valeurs isolées dans un intervalle donné, ces valeurs sont souvent des valeurs entiers. Exemple : nombre d'enfant : 1,…6 ; ou encore le nombre de matières à étudiés : 1 ;….15.
   - Caractères quantitatifs continus : c'est un caractère qui ne peut prendre que des valeurs qui appartiennent à un intervalle donné. Exemple : âge ; taille ; poids ; revenu…

Les caractères sont aussi appelés VARIABLES STATISTIQUES donc on distingue trois types de variables à savoir : variables qualitatives, variables quantitatives discrètes et variables quantitatives continues.

# II-TABLEAUX ET GRAPHIQUES :

L'information statistique collectée sous forme de données individuelles, n'est pas facilement exploitable et sa manipulation est lourde, il est donc nécessaire de résumer les caractères sous forme de tableaux.

Par ailleurs, l'information statistique ne peut jamais être publiée sous sa forme brute, il faut la représenter sous forme simplifiée par des tableaux : on parle alors de données groupées ou classées ce qui nous donne ce qu'on appelle **distribution statistique (DS).**

Une DS est une répartition de la population observée selon les différentes modalités du ou des caractères retenus. Si on retient un seul caractère, alors la DS est dite à une seule dimension et on présente alors un tableau à une seule dimension ou encore un *tableau à simple entrée.*

La DS peut aussi être représentée par un graphique qui a l'avantage de donner une lecture visuelle immédiate des aspects dominants.

Les tableaux et les graphiques diffèrent selon la nature du caractère étudié.

# Section 1 : REPRESENTATION DES CARACTERES QUALITATIFS

**1)** tableaux statistiques :

Soit une population P de **N** individus, sur laquelle on observe le caractère qualitatif **C** qui comporte **k** modalités (nomenclatures) notés $M_1$ ; $M_2$ ......... $M_k$. Soit $n_i$ le nombre d'individus de la population qui présentent la modalité $M_i$ du caractère C. $n_i$ est l'effectif de la modalité $M_i$ et on a $\sum_{i=1}^{k} n_i = N$.

On appelle tableau statistique de la population P décrite selon le caractère C, le tableau des couples ($M_i$ , $n_i$).

Répartition de la population P selon le caractère C :

| **modalités** | **effectifs** |
|---|---|
| $M_1$ | $n_1$ |
| ... | ... |
| $M_i$ | $n_i$ |
| ... | ... |
| $M_k$ | $n_k$ |
| **total** | **N** |

Pour représenter un tableau statistique, trois règles doivent être respectées :

- il faut mettre le titre du tableau ;
- il faut mettre la source ou l'origine de l'information chiffrées (en bas et à droite) ;
- il faut indiquer les unités utilisées pour les effectifs.

On appelle fréquence de la modalité $M_i$ le quotient : $f_i = \frac{n_i}{N}$ et on peut aussi définir le pourcentage de la manière suivante : $p_i = 100 \times f_i$. Ainsi : $\sum_{i=1}^{k} f_i = 1 \; et \; \sum_{i=1}^{k} p_i = 100$.

a) **Exemple :** répartition des 1000 voitures du gouvernorat de Nabeul selon la couleur :

| **couleurs** | **effectifs** | **fréquences** |
|---|---|---|
| Bleu | 150 | ……. |
| Noire | 280 | ……. |
| Rouge | 220 | ……. |
| Jaune | 250 | ……. |
| autres | 100 | ……. |
| **total** | **1000** | **1** |

Source : exemple

**P** : Population étudiée :
**C** : Caractère :
**M** : Ensemble des modalités :

**2)** représentation graphique :

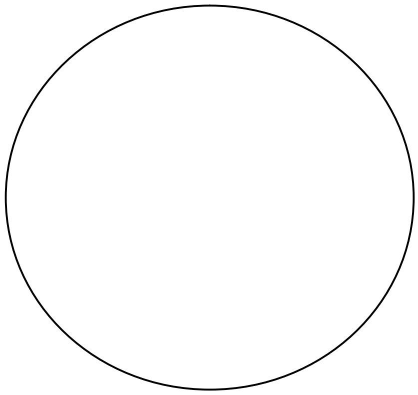
Pour les caractères qualitatifs, deux types de graphiques sont utilisés :

a) **Diagramme en secteurs :**

La distribution est représentée par un cercle divisé en **k** secteurs (chaque modalité sera représentée à l'aide d'un secteur sur le cercle), la superficie du secteur (l'angle de chaque secteur noté $\alpha_i$) est proportionnelle à l'effectif ou à la fréquence de cette modalité. On a donc $\alpha_i = 360° f_i$

| **couleur** | **angle $\alpha_i$** | **pourcentage $p_i$** |
|---|---|---|
| Bleu | ……. | ……. |
| Noire | ……. | ……. |
| Rouge | ……. | ……. |
| Jaune | ……. | ……. |
| autres | ……. | ……. |
| **total** | **360** | **100** |

**Distribution des voitures selon la Couleur**



b) **Le graphique en tuyaux d'orgue :**

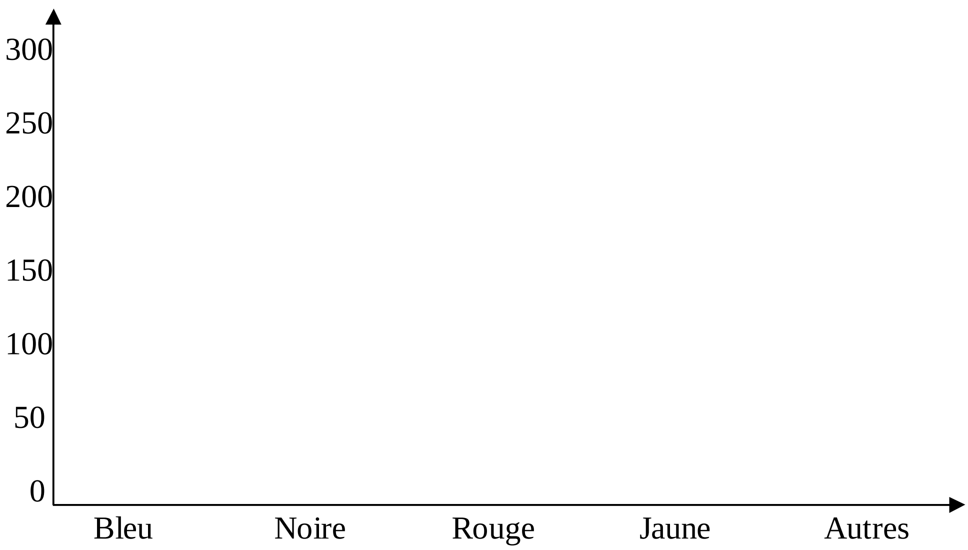
C'est une représentation graphique d'un ensemble de rectangles dans un repère orthonormé ayant :

En abscisses : les modalités du caractère. La largeur de chaque rectangle est la même quelle que soit la modalité, la largeur n'est pas une mesure.

En ordonnées : les valeurs des effectifs ou des fréquences.

La représentation des rectangles peut se faire selon un ordre arbitraire des modalités.

**Distribution des voitures selon la Couleur**



L'aire ou la surface du rectangle est proportionnelle à l'effectif ou à la fréquence de la modalité puisque les rectangles ont tous la même largeur (choisie par l'utilisateur).

## Section 2 : **REPRESENTATION DES CARACTERES QUANTITATIFS DISCRETS**

**1)** Les tableaux :

La présentation générale sous forme de tableau se présente comme suit :

Répartition de la population P selon le caractère C :

| **modalités** | **effectifs** | **fréquences** |
|---|---|---|
| $M_1$ | $n_1$ | $f_1$ |
| ... | ... | ... |
| $M_i$ | $n_i$ | $f_i$ |
| ... | ... | ... |
| $M_k$ | $n_k$ | $f_k$ |
| **total** | **N** | **1** |

Les $M_i$ sont la plupart des cas des nombres entiers (variable discrète). La dernière modalité $M_k$ est souvent un regroupement de plusieurs valeurs dont les effectifs sont trop faibles pour constituer des modalités à part entière.

a) **Exemple :** répartition des étudiants d'un amphi selon le nombre de frères et sœurs :

| **nombre de frères et sœurs** | **effectifs** | **fréquences** |
|---|---|---|
| 0 | 18 | ……. |
| 1 | 30 | ……. |
| 2 | 60 | ……. |
| 3+ | 42 | ……. |
| **total** | **150** | **1** |

**2)** fonction de répartition : fréquences cumulés croissants :

Les caractères quantitatifs discrets ont des modalités ordonnées (pas comme le cas des caractères qualitatifs), de ce fait, on peut construire la fonction de répartition.

Notation : X est la valeur du caractère quantitatif discret ;
x est une valeur particulière donnée à ce caractère.

a) **Définition :**

La fonction de répartition d'un caractère quantitatif discret est une application F de IR dans l'intervalle [0,1] définie de la façon suivante :

$$F : IR \longrightarrow [0,1]$$
$$x \longrightarrow F(x) = P\,(X<x)$$

De cette définition de la fonction de répartition découlent les considérations suivantes :

- F est définie quelle que soit x appartenant à IR, x correspond ou non à une modalité de X ;
- $F(x) = 0$ quelle que soit $x \leq M_1$ en particulier $F\,(M_1) = 0$ ;
- $F\,(M_2) = P\,(X< M_2) = P\,(X = M_1) = f_1$ ;
- $F\,(M_3) = P\,(X< M_3) = P\,(X = M_1$ ou $X = M_2\,) = f_1 + f_2$ ;
- En générale pour $i \geq 2$ ;

$F(M_i) = P\,(X< M_i) = P\,(X = M_1$ ou $X = M_2$ ..... ou $X = M_{i-1}) = f_1 + f_2 + .... + f_{i-1}$;

- $F(x) = 1$ quelle que soit $x > M_k$ ;
- F est une fonction non décroissante, elle est croissante ou constante puisque ces valeurs sont de plus en plus grandes ou constantes. D'une manière générale, la fonction de répartition d'un caractère discret est constante par intervalle.

Tableau des valeurs remarquables de la fonction de répartition :

| $M_i$ | $f_i$ | $F(M_i)$ | $N(M_i)$ |
|---|---|---|---|
| $M_1$ | $f_1$ | 0 | 0 |
| $M_2$ | $f_2$ | $f_1$ | $n_1$ |
| .... | .... | .... | .... |
| $M_i$ | $f_i$ | $f_1 + f_2 + .... + f_{i-1}$ | $n_1 + n_2 + .... + n_{i-1}$ |
| .... | .... | .... | .... |
| $M_k$ | $f_k$ | $f_1 + f_2 + .... + f_i .... + f_{k-1}$ | $n_1 + n_2 + .... + n_i .... + n_{k-1}$ |

- la grandeur F ($M_i$) est appelée fréquence cumulée croissante de la modalité $M_i$ ;
- On peut aussi définir des effectifs cumulés croissants avec $N(x) = E(X < x)$
- On peut définir des fréquences cumulés décroissants avec $G(x) = P(X \geq x) = 1 - F(x)$. C'est la répartition d'individus ayant une valeur de la variable supérieur ou égale à $x$.

b) **Application :**

Reprenant l'exemple précédent : de la répartition des étudiants d'un amphi selon le nombre de frères et sœurs :

| **Nombre de frères et sœurs** | $n_i$ | **$f_i$** | **N(x)** | **F(x)** | **G(x)** |
|---|---|---|---|---|---|
| 0 | 18 | 0,12 | ……. | ……. | ……. |
| 1 | 30 | 0,20 | ……. | ……. | ……. |
| 2 | 60 | 0,40 | ……. | ……. | ……. |
| 3+ | 42 | 0,28 | ……. | ……. | ……. |
| **total** | **150** | 1 | --- | --- | --- |

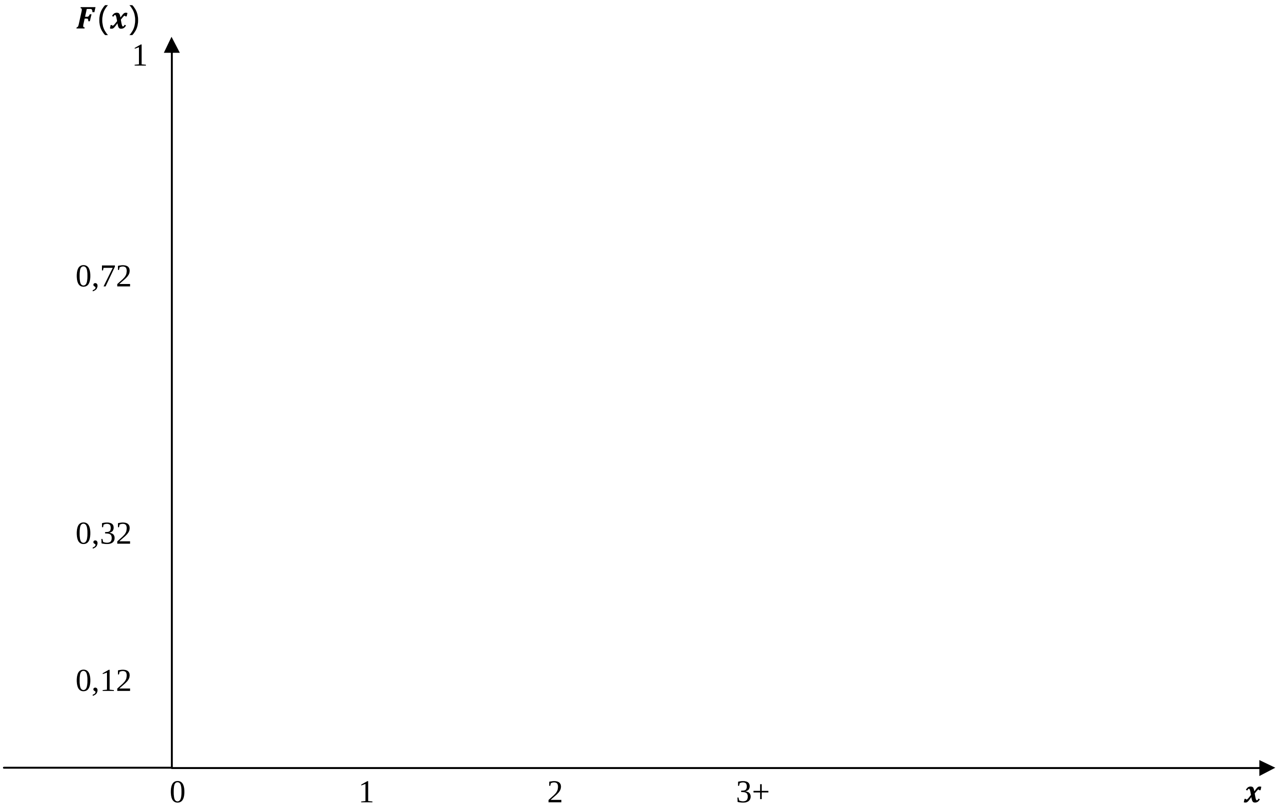
**3)** Les graphiques :

Les caractères quantitatifs discrets donnent lieu à deux types de représentations graphiques.

a) **Diagramme intégral ou courbe cumulative :**

C'est la courbe de la fonction de répartition, qui est une courbe en escalier et discontinue.

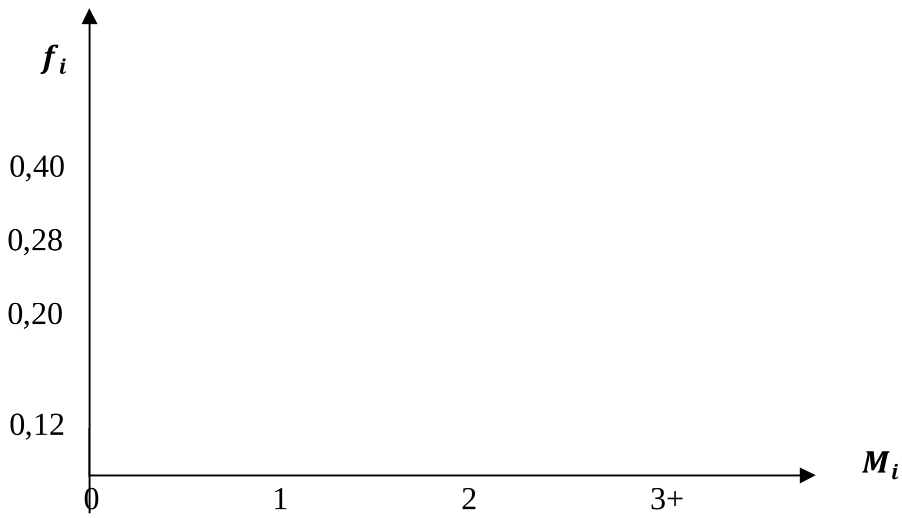
Représentation graphique de la fonction de répartition de la variable nombre de frères et sœurs :




b) **Diagramme différentiel (diagramme en bâtons)**

Chaque modalité est représentée (dans un repère orthonormé : modalité en abscisses et effectifs ou fréquences en ordonnés) par un segment vertical dont la longueur est proportionnelle à l'effectif ou à la fréquence.

- Reprenant le même exemple :




## Section 3 : **REPRESENTATION DES CARACTERES QUANTITATIFS CONTINUS**

**1)** Modalités et tableaux :

Les modalités appartiennent à des intervalles [a, b] de IR avec a et b qui sont respectivement la plus petite et la plus grande valeur observées.

[a, b] sera donc subdivisé en **k** sous – intervalles disjoints $[e_{i-1}, e_i[$ : par convention fermé à gauche et ouvert à droite. Chaque sous – intervalle est appelé classe, la différence $e_i - e_{i-1} = a_i$ s'appelle amplitude de la classe.
On pose $a = e_0$ et $b = e_k$
On définit ainsi k modalités $M_i$ constituées par k classes :
$M_1 = [e_0, e_1[$ ; $M_2 = [e_1, e_2[$ ; ................ $M_i = [e_{i-1}, e_i[$ ; ......... $M_k = [e_{k-1}, e_k[$ ;
L'effectif $n_i$ de la classe $M_i = [e_{i-1}, e_i[$est le nombre d'individus qui ont une valeur de la variable supérieur ou égale à $e_{i-1}$ et strictement inférieur à $e_i$.
Le tableau se présente alors de la manière suivante :

| **modalités** | **effectifs** | **fréquences** |
|---|---|---|
| $M_1$ | $n_1$ | $f_1$ |
| $M_2$ | $n_2$ | $f_2$ |
| ... | ... | ... |
| $M_i$ | $n_i$ | $f_i$ |
| ... | ... | ... |
| $M_k$ | $n_k$ | $f_k$ |
| **total** | **n** | **1** |

a) **Exemple :** répartition des ouvriers d'une usine selon les dépenses journalière :

| **Dépenses en DT** | **effectifs** | **fréquences** |
|---|---|---|
| [10, 15[ | 42 | ……. |
| [15, 18[ | 66 | ……. |
| [18, 24[ | 103 | ……. |
| 24 et + | 89 | ……. |
| **total** | **300** | **1** |

b) **Remarque :**
***R1 :*** il arrive souvent que les bornes $e_0$ et $e_k$ ne soient pas définies avec précision c'est à dire $M_1$ : moins que $e_0$ et $M_k$ : plus que $e_k$.
Cette imprécision provient du fait que les valeurs limites $e_0$ et $e_k$ sont assez éloignées. Alors par convention : pour fixer $e_0$, on prend la même amplitude que la classe suivante, de même pour fixer $e_k$ on prend l'amplitude de la classe précédente.
***R2 :*** le problème du découpage en classe : autrement dit le problème du choix du nombre k, il faut retenir qu'il n'existe pas de règles universelles et que le nombre de classes ne doit pas être ni trop grand ni trop petit. En général, k appartient à l'intervalle] 6, 13[avec des effectifs plus ou moins équilibrés entre les classes.

**2)** <u>La fonction de répartition *:*</u>
On garde la même définition que dans le cas discret. Cependant certaines remarques doivent êtres faites :
$F(x) = 0$ quelle que soit $x \leq e_0$ c'est à dire $F(e_0) = 0$ ;
$F(e_1) = f_1$ ;
$F(e_2) = f_1 + f_2$ ;
........
$F(e_i) = f_1 + f_2 + ...... + f_i$;
........
$F(e_k) = f_1 + f_2 + ...... + f_i + ....... + f_k = 1$;

Quelle que soit x appartenant à $[e_{i-1}, e_i[$ d'une classe donnée, F(x) n'est pas déterminer avec exactitude. On supposera que la fonction F est linéaire et non pas constante entre $e_{i-1}$ et $e_i$ comme dans le cas discret. Avec cette hypothèse de linéarité de la fonction de répartition à l'intérieur des classes, on peut déterminer F(x) quelle que soit x, cette détermination se fait par **interpolation linéaire**.

Interpolation linéaire:

$e_{i-1} \longrightarrow F(e_{i-1})$

$x \longrightarrow F(x)$

$e_i \longrightarrow F(e_i)$

$$\frac{F(x)-F(e_{i-1})}{x-e_{i-1}} = \frac{F(e_i)-F(e_{i-1})}{e_i-e_{i-1}}$$

a) **Exemple :** Reprenant l'exemple des dépenses journalières :

| **Dépenses en DT** | **fréquences** | **fonction de répartition** |
|---|---|---|
| [10, 15[ | 0,140 | ……. |
| [15, 18[ | 0,220 | ……. |
| [18, 24[ | 0,343 | ……. |
| [24, 30[ | 0,297 | ……. |
| **total** | **1** | **----** |

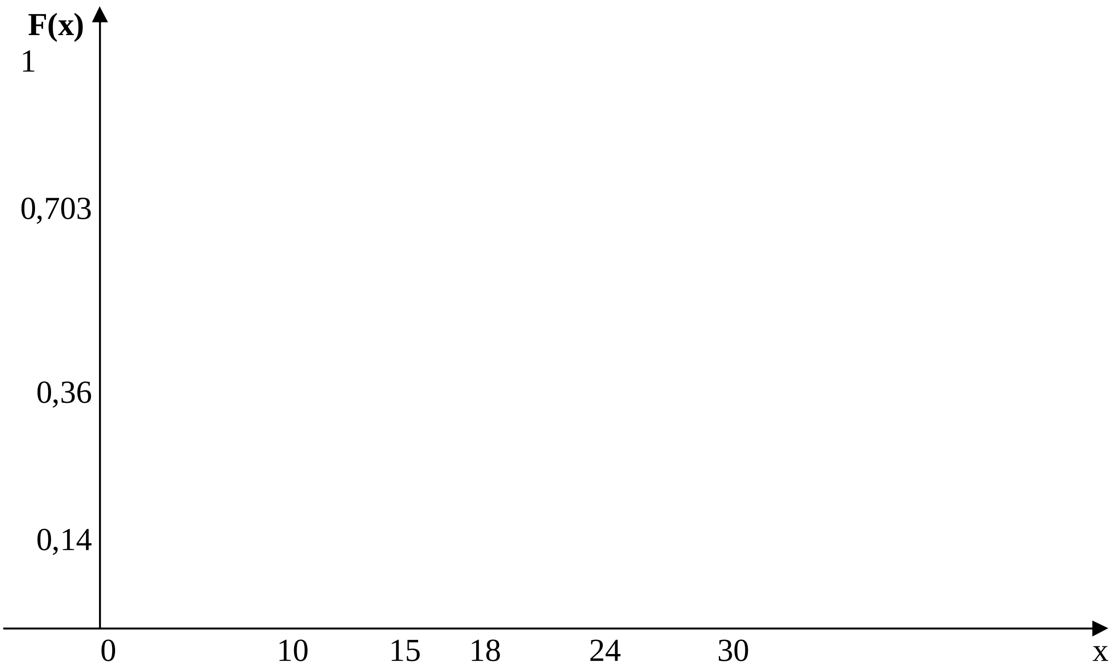
**3)** Les graphiques *:*

Les caractères quantitatifs continus font l'objet de deux représentations graphiques :

a) **Diagramme intégral ou courbe cumulative :**

C'est la courbe de la fonction de répartition qui est une courbe croissante et continue.

Représentation graphique de cette fonction de répartition :




b) **Histogramme ou diagramme différentiel :**

C'est un ensemble de k rectangles superposés (l'un à coté de l'autre), il s'agit d'un rectangle par modalité ou par classe. La base ou la largeur de chaque rectangle est égale à l'amplitude de la classe et la longueur est égale à son effctif ou à sa fréquence. Deux cas peuvent se présenter :

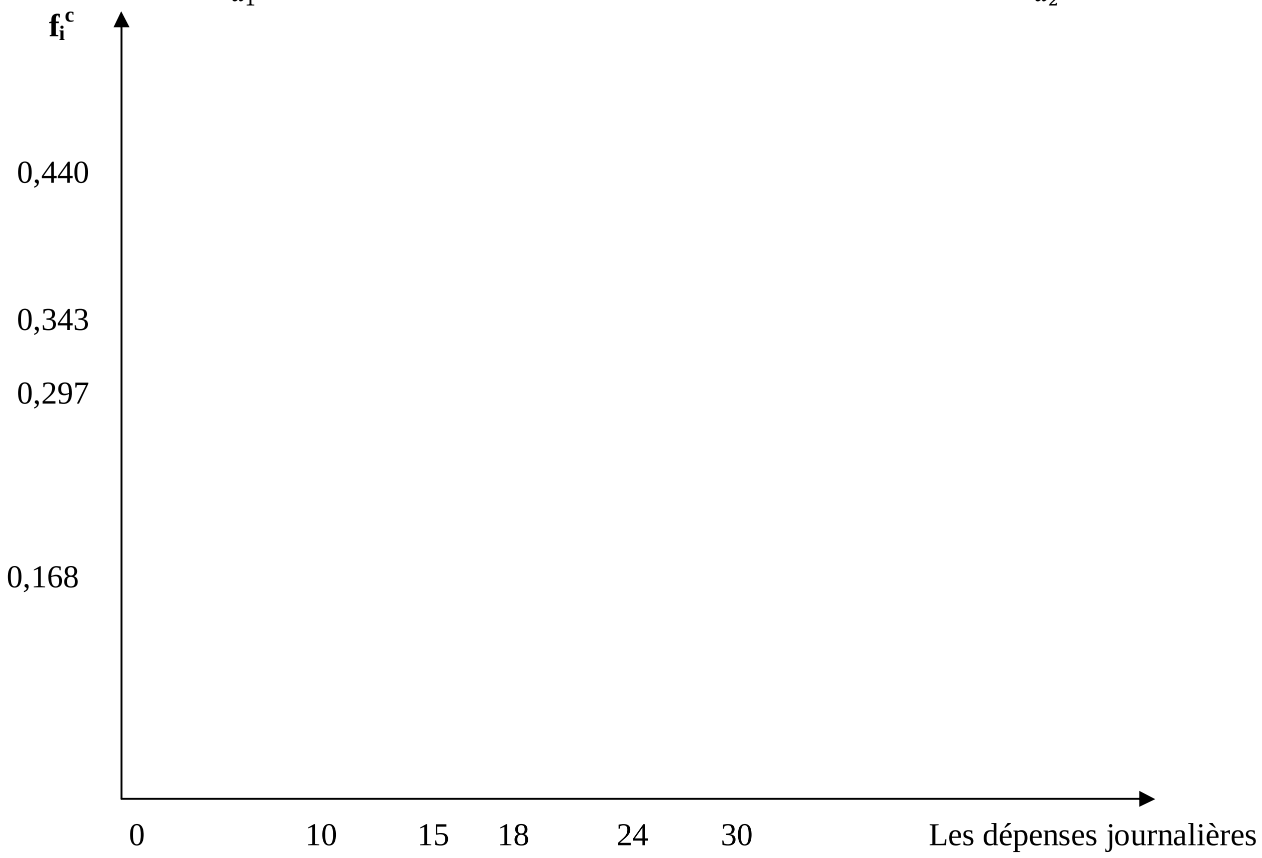
**Cas n°1 :** toutes les amplitudes sont égales : la représentation de l'histogramme est directe.
**Cas n°2 :** les classes ont des amplitudes différentes, on choisit par exemple la plus petite amplitude ou la plus fréquente comme amplitude de reférence et on calcule les fréquences corrigées $f_i^c$ ou les effectifs corrigées $n_i^c$qui vont représentées les longueurs des rectangles.
$f_i^c = f_i \frac{a_{ref}}{a_i}$ $ou\ bien$ $n_i^c = n_i \frac{a_{ref}}{a_i}$.avec $a_{ref}$ $et$ $a_i$ sont respectivement l'amplitude de référence et de la classe.
En résumé, pour construire un histogramme, quatre étapes sont nécessaires :

- calculer l'amplitude de chaque classe ;
- choisir l'amplitude de reférence qui est la plus petite ou la plus répandue ;
- calculer les fréquences corrigées ;
- tracer l'histogramme qui est un ensemble de k rectangles juxtaposés dont la largeur est égal à l'amplitude initiale de la classe et la longueur égale à la fréquence corrigée.

Reprenant une autre fois le même exemple des dépenses journalières :
Choisissons 6 comme amplitude de référence.

| Dépenses en DT | fréquences | Les Amplitudes | Fréquences corrigées |
|---|---|---|---|
| [10, 15[ | 0,140 | ……. | ……. |
| [15, 18[ | 0,220 | ……. | ……. |
| [18, 24[ | 0,343 | ……. | ……. |
| [24, 30[ | 0,297 | ……. | ……. |
| **total** | **1** | **----** | **-------** |

$f_1^c = f_1 \times \frac{a_{ref}}{a_1} =$ ; $f_2^c = f_2 \times \frac{a_{ref}}{a_2} =$



Chapitre 2

# LES PARAMETRES DE TENDANCE CENTRALE ET DE DISPERSION

Les tableaux et graphiques examinés au cours du chapitre précédent rendent l'information statistique brute plus lisible et plus manipulable.
Ce chapitre a pour objectif de résumer encore davantage la DS par quelques grandeurs significatives appelées paramètres de tendance centrale et de dispersion.

## I- LES PARAMETRES DE TENDANCES CENTRALE OU DE POSITION :

### Section 1 : **LE MODE : Mo**

1) Définition :

- Le mode Mo est la valeur ou l'état de la variable statistique c'est à dire la modalité qui correspond à l'effectif le plus élevé ou à la fréquence la plus élevée.
C'est donc la valeur la plus fréquente pour la population.
- Le mode Mo est la valeur la plus fréquente dans une série statistique.

2) Détermination :

- Dans les cas qualitatifs et quantitatifs discrets, la détermination du Mode est immédiate ; exemples :

a) **Le cas quantitatifs discrets :** La répartition du nombre d'enfants par familles :

| Nombre d'enfants | Nombre de familles |
|---|---|
| 0 | 15 |
| 1 | 25 |
| 2 | 10 |
| 3 | 35 |
| **4** | **74** |
| $5^+$ | 12 |

Mo =

b) **Le cas qualitatif :** La répartition des voitures selon leurs couleurs :

| Couleurs de voitures | Nombre de voitures |
|---|---|
| Noire | 17 |
| **Jaune** | **59** |
| Bleu | 53 |
| Autre | **59** |

Mo =

c) **Le cas d'un caractère quantitatif continu** : Il s'agit de déterminer une classe modale, deux cas sont possibles :

❖ Toutes les classes ont la même amplitude : la classe modale est celle qui correspond à l'effectif le plus élevé ou à la fréquence la plus élevée

- les classes ont des amplitudes différentes : on calcule les fréquences corrigées, la classe modale sera celle qui correspond à la fréquence corrigée la plus élevée.

**Remarque :**

***R1 :*** sur l'histogramme, la classe modale est celle qui correspond au rectangle ayant la longueur la plus élevée ;

***R2 :*** on peut considérer le centre $C_i$ d'une classe modale comme étant le mode Mo ;

***R3 :*** une DS peut être PLURIMODALE ; si elle admet plus qu'un mode ou UNIMODALE ; si elle admet un mode unique.

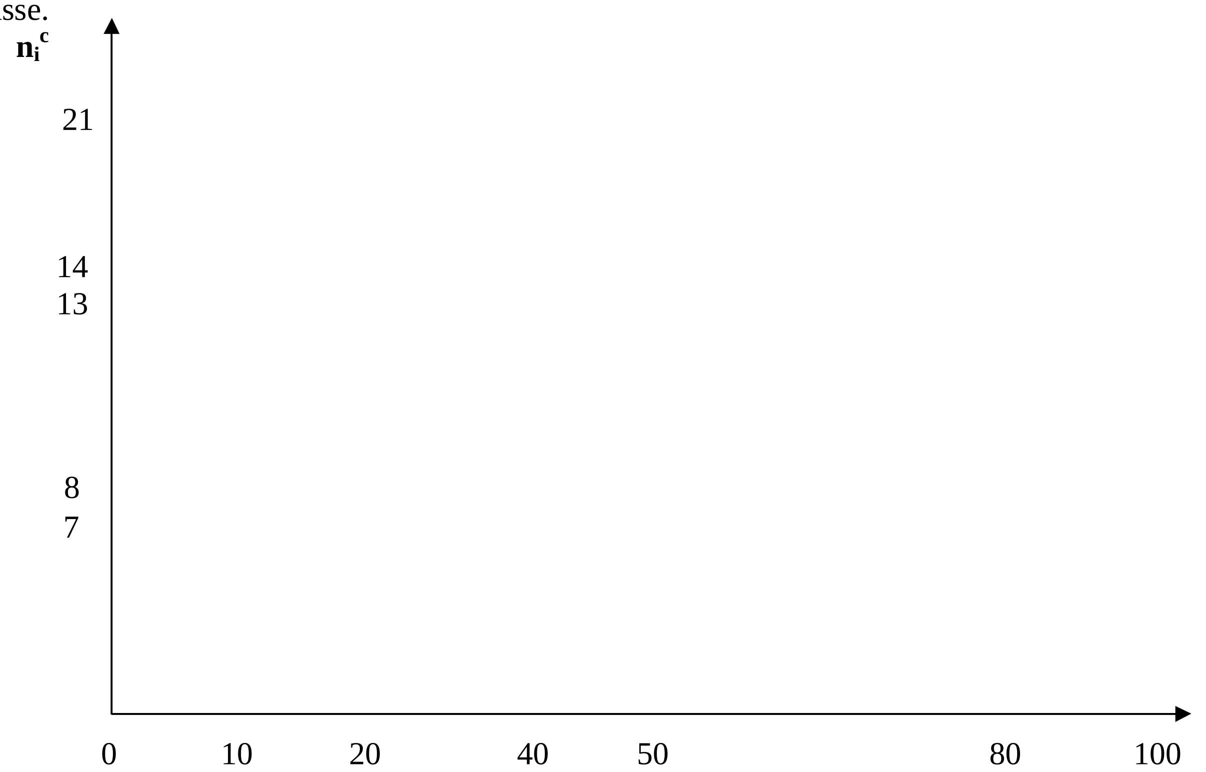
- Détermination graphique et par le calcul:

Une entreprise a enquêté 100 délégués médicaux sur le nombre de Km qu'ils effectuaient par jour pour la vente de ces produits. Le tableau suivant résume les résultats obtenus.

| **Trajet** | **Effectifs** $n_i$ | **Amplitude** $a_i$ | **Effectifs corrigés** $n_i^c$ | **fréquences** $f_i$ | **fréquences corrigées** $f_i^c$ |
|---|---|---|---|---|---|
| [10, 20[ | 13 | ……. | ……. | ……. | ……. |
| [20, 40[ | 28 | ……. | ……. | ……. | ……. |
| [40, 50[ | 21 | ……. | ……. | ……. | ……. |
| [50, 80[ | 24 | ……. | ……. | ……. | ……. |
| [80, 100[ | 14 | ……. | ……. | ……. | ……. |
| **total** | **100** | **----** | **----** | **1** | **----** |

Choisissons 10 comme amplitude de référence.

$f_i^c = f_i \frac{a_{ref}}{a_i}$ $ou\ bien$ $n_i^c = n_i \frac{a_{ref}}{a_i}$.avec $a_{ref}$ $et$ $a_i$ sont respectivement l'amplitude de référence et de la classe.




Mo =

$$Mo = L_{Mo} + \frac{d_1}{d_1 + d_2} \times A$$

Avec :
- $L_{Mo}$ : La limite (borne) inférieure de la classe modale,
- $d_1$ : La fréquence corrigée de la classe modale - la fréquence corrigée de la classe précédente,
- $d_2$ : La fréquence corrigée de la classe modale - la fréquence corrigée de la classe suivante,
- A : L'amplitude de la classe modale.

d) **Application :** $Mo \in [40{,}50[$

$$Mo = \cdots + \frac{\ldots \quad - \quad \ldots}{(\ \ldots - \ldots\ ) + (\ \ldots - \ldots\ )} \times \ = \ \ldots + \frac{\ldots \quad - \quad \ldots}{(\ \ldots - \ldots\ ) + (\ \ldots - \ldots\ )} \times \ \ldots = \cdots \quad \text{Km/j.}$$

Le kilométrage le plus souvent effectué par les délégués médicaux est de 43 Km et demi par jour (la plupart des délégués médicaux ont effectué 43,5 Km/j).

## Section 2 : **LA MEDIANE : Me**

**1)** Définition :

La médiane Me est la valeur de la variable telle que 50 % de l'effectif total ait une valeur inférieur strictement à Me et 50 % ait une valeur supérieure ou égale à Me. La médiane partage alors la population en deux parties d'effectifs égaux.
Analytiquement et à partir de la fonction de répartition, la médiane Me est définie comme suit : $F(Me) = 0{,}5$

a) **Le cas qualitatif** : Dans ce cas, les modalités ne sont pas des valeurs ordonnées donc pas de fonction de répartition et par la suite pas de médiane.

b) **Le cas discret** : Dans ce cas, en général la médiane n'existe pas, cependant dans certains cas rares, on tombe exactement sur une modalité $M_i$ qui correspond à une valeur 0,5 de la fonction de répartition et donc $M_i$ sera la médiane. On peut prendre comme médiane la modalité qui correspond à une valeur de la fonction de répartition proche de 0,5 c'est à dire appartenant à l'intervalle [0,48 ; 0,52]. Sinon la médiane n'existe pas.

c) **Le cas continu** : Dans ce cas, deux étapes sont nécessaires pour déterminer la médiane :
* déterminer la classe médiane $[e_{i-1} ; e_i]$ en vérifiant : $F(e_{i-1}) < 0{,}5 < F(e_i)$
* on déterminera la valeur exacte de la médiane par interpolation linéaire à l'intérieur de la classe médiane.

➢ Exemple : Reprenant l'exemple précédent :

| **Trajet** | **effectifs** | **fréquences** | **fréquences cumulées croissantes F(X)** | **fréquences cumulées décroissantes G(X)** |
|---|---|---|---|---|
| [10, 20[ | 13 | 0,13 | ……. | ……. |
| [20, 40[ | 28 | 0,28 | ……. | ……. |
| [40, 50[ | 21 | 0,21 | ……. | ……. |
| [50, 80[ | 24 | 0,24 | ……. | ……. |
| [80, 100[ | 14 | 0,14 | …..... | ……. |
| **total** | **100** | **1** | **---** | **----** |

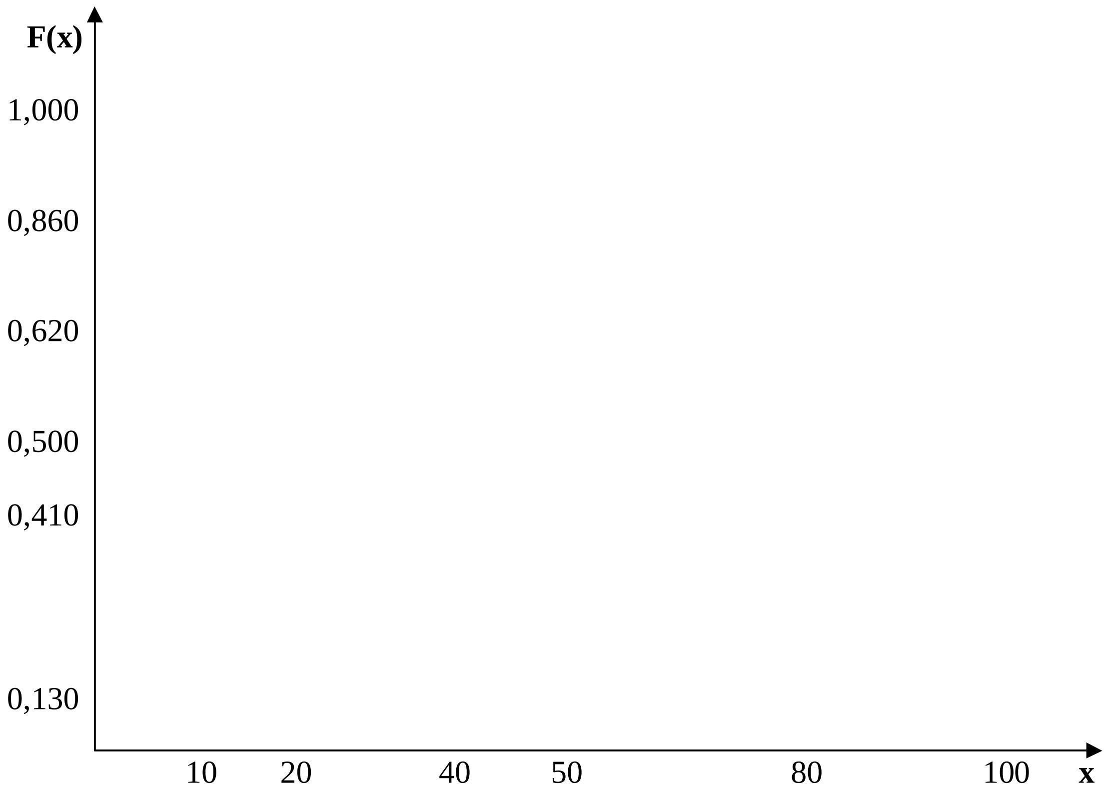
On a donc comme classe médiane l'intervalle [ .... ; .... [ $\Rightarrow \begin{cases} F(40) = 0,\ldots & < 0,5 \\ F(50) = 0,\ \ldots. & > 0,5 \end{cases}$

… ----------------- …
Me ----------------- **0,5**
… ------------------ …

À l'aide d'une interpolation linéaire, on aura Me = 44,286 Km/j.
50% des représentants ont effectué un trajet inférieur à 44,286 Km/j et l'autre 50% ont effectué plus de 44,286 Km/j dans leur déplacement.

**2)** Détermination graphique de la médiane :

En utilisant la courbe de la fonction de répartition (courbe cumulative), on prend le point d'ordonnée 0,5 ; on le projette sur la courbe et ensuite sur l'axe des abscisses ce qui nous permet d'obtenir la Médiane.
Représentation graphique de cette fonction de répartition :




Me = 44,286 Km/j

## Section 3 : **La MOYENNE ARITHMETIQUE :** $\overline{X}$

**1)** définition :

La moyenne arithmétique d'une variable statistique relative à une population est la somme des valeurs observées divisée par le nombre d'observations. Pour la suite de ce cours, on convient de choisir les notations suivantes :

- pour le cas des variables aussi bien discrètes que continues, on adopte la notation suivante : $x_1$, $x_2$, $x_3$, ......$x_i$, .......... $x_k$ pour les modalités et $n_1$, $n_2$, $n_3$, ......$n_i$, .......... $n_k$ pour les effectifs qui représentent les pondérations.

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

Pour le cas discret, $x_i$ est la valeur de la modalité. Dans le cas continu, on prend $c_i$ le centre de la classe au lieu de $x_i$ puisque les modalités sont des intervalles et non pas des valeurs isolées.

Il faut remarquer que la moyenne arithmétique n'est pas une donnée qui a une existence réelle, qu'elle n'est pas un nombre entier. Même dans le cas discret, si la variable est par exemple le nombre d'enfant, on peut accepter une moyenne arithmétique de 2,58.

a) **Exemples :**

**i.** Soit la répartition des étudiants d'une classe de TD selon le nombre de leurs puces téléphoniques :

| **Nombre de puce ($x_i$)** | **Nombre d'étudiants ($n_i$)** |
|---|---|
| 1 | 8 |
| 2 | 14 |
| 3+ | 3 |
| **Total** | **25** |

$$\overline{X} = \frac{(\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ )}{25} = \dots$$

Le nombre moyen de puces téléphoniques de ces étudiants est de 1,8.

**ii.** Reprenant l'exemple des représentants :

| **Trajet** | **effectifs** | **fréquences** | **Centre des classes $c_i$** |
|---|---|---|---|
| [10, 20[ | 13 | 0,13 | ……. |
| [20, 40[ | 28 | 0,28 | ……. |
| [40, 50[ | 21 | 0,21 | ……. |
| [50, 80[ | 24 | 0,24 | ……. |
| [80, 100[ | 14 | 0,14 | ……. |
| **total** | **100** | **1** | **---** |

$$\overline{X} = \frac{(\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ )}{100} = \mathbf{48\ Km/j}$$

$$\bar{X} = (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) = \mathbf{48}\ \boldsymbol{KM/j}$$

Les représentants effectuent en moyenne un trajet de 48 Km/j.

b) **Propriétés :**

**$P_1$-** propriété de linéarité :

Si la variable Y est une transformation linéaire d'une autre variable X, alors la même transformation linéaire est valable sur leurs moyennes arithmétiques : $Y = aX + b$ alors $\bar{Y} = a\bar{X} + b$

**$P_2$-** si $P_1$ et $P_2$ sont deux sous populations d'effectifs respectifs $N_1$ et $N_2$ avec des moyennes arithmétiques respectives $\overline{X_1}$ et $\overline{X_2}$ alors la population $P = P_1\ U\ P_2$ d'effectifs total : $N = N_1 + N_2$ a pour moyenne arithmétique :

$$\bar{X} = \frac{N_1\,\overline{X_1} + N_2\overline{X_2}}{N}$$

Cette propriété reste valable pour le cas de plusieurs sous populations.

## Section 4 : **AUTRES MOYENNES ET GENERALISATION**

**1)** La moyenne quadratique : Noté Q est celle dont le carré est égal à la moyenne arithmétique des carrés des valeurs de la variable prises par la population :

$$Q^2 = \frac{\sum_{i=1}^{k} n_i x_i^2}{N} \quad \text{et donc } Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{N}}$$

**Pour le cas continu, $x_i$ sera remplacé par $c_i$.**

a) **Application :** Reprenant l'exemple de la répartition des étudiants d'une classe de TD :

$$Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{N}} = \sqrt{\frac{(\ \dots \times \dots\ ) + (\ \dots \times \dots\ ) + (\ \dots \times \dots\ )}{\mathbf{25}}} = \mathbf{1{,}908}$$

**2)** La moyenne géométrique : Noté G est la valeur de la variable telle que le logarithme de G soit égale à la moyenne arithmétique des logarithmes des valeurs observées :

$$Log\ G = \frac{\sum_{i=1}^{k} n_i \times Log\ (x_i)}{N} \quad \text{Et donc} \quad G = \sqrt[N]{{x_1}^{n_1}{x_2}^{n_2} \dots {x_i}^{n_i} \dots {x_k}^{n_k}}$$

a) **Application :** Reprenant l'exemple de la répartition des étudiants d'une classe de TD :

$$G = \sqrt[N]{{x_1}^{n_1}{x_2}^{n_2} \dots {x_i}^{n_i} \dots {x_k}^{n_k}} = \sqrt[25]{\dots \times \dots \times \dots} = 1{,}682$$

b) **Propriétés :**

- La moyenne géométrique du produit est égale au produit des moyennes géométriques :

$$Z_i = X_i \times Y_i \text{ alors } G(Z) = G(X) \times G(Y)$$

- La moyenne géométrique du quotient est égale au quotient des moyennes géométriques :

$$Z_i = \frac{X_i}{Y_i} \text{ alors } G(Z) = \frac{G(X)}{G(Y)}.$$

**3)** La moyenne harmonique : Noté H est la valeur de la variable dont l'inverse est égale à la moyenne arithmétique des inverses des observations :

$$\frac{1}{H} = \frac{1}{N}\sum_{i=1}^{k}\frac{n_i}{x_i} \quad \text{Et donc} \quad H = \frac{N}{\sum_{i=1}^{k}\frac{n_i}{x_i}}$$

a) **Application :** Reprenant l'exemple de la répartition des étudiants d'une classe de TD :

$$H = \frac{N}{\sum_{i=1}^{k}\frac{n_i}{x_i}} = \frac{25}{\frac{\dots}{\dots} + \frac{\dots}{\dots} + \frac{\dots}{\dots}} = 1{,}563$$

Remarque :

- La moyenne quadratique : est utilisée surtout en économétrie.
- La moyenne géométrique est utilisée pour calculer le taux de croissance moyen ainsi que les anuitées.
- La moyenne harmonique est utilisée pour mesurer les unités composées comme par exemple : Km/h ; DT/$...

-Dans le cas continu et lors du calcul de ces différentes moyennes, $x_i$ sera remplacé par $c_i$.

**4)** Comparaisons des différentes moyennes : pour une même distribution ou variable, les différentes moyennes vérifient l'inégalité suivante : $\boldsymbol{H} < G < \overline{\boldsymbol{X}} < Q$

# II- LES PARAMETRES DE DISPERSION ET DE FORME :

Les indicateurs de tendance centrale (Mode, Médiane et moyenne arithmétique) sont insuffisants pour permettre de résumer et de comparer les séries statistiques pour la simple raison qu'on peut avoir plusieurs séries qui possèdent les mêmes valeurs pour ces paramètres mais avec des distributions qui se font d'une manière nettement différentes, d'où la nécessité de calculer d'autres indicateurs capable de rendre compte des écarts entre les différentes valeurs observées et les valeurs centrales.

## Section 1 : L'ETENDUE : e

On appelle étendue d'une série statistique, la différence entre la plus élevée et la plus faible des valeurs observées, soit :

$$e = x_{max} - x_{min}$$

Dans l'exemple précèdent $e = \cdots - \cdots = \cdots$

## Section 2 : LES ECARTS INTERQUANTILES : LES QUANTILES

Comme pour la médiane, on s'intéresse ici aux valeurs de la variable qui partagent la population en 4, en 10 et 100 parties de même effectif. Ces valeurs sont appelées respectivement quartiles, déciles et centiles.

**1)** Les quartiles :

Ils sont en nombre de trois notés $Q_1$, $Q_2$ et $Q_3$, ils partagent la population en 4 groupes d'effectif égal.
**$Q_1$** : premier quartile : valeur de la variable telle que 25 % des observations lui soient inférieures.
$F(Q_1) = 0,25$ ;
**$Q_2$** : deuxième quartile : valeur de la variable telle que 50 % des observations lui soient inférieures.
$F(Q_2) = 0,5$ ($Q_2 = Me$) ;
**$Q_3$** : troisième quartile : valeur de la variable telle que 75 % des observations lui soient inférieures.
$F(Q_3) = 0,75$ ;
L'intervalle d'amplitude $Q_1$- $Q_3$ est appelé intervalle interquartile et noté : $[Q_1,Q_3]$.
Cet intervalle comporte 50% des observations.

**2)** Les déciles :

En nombre de neuf notés $D_1$, $D_2$, ....., $D_9$, ils partagent la population en 10 groupes d'effectif égal.

**a) Exemple :**

**$D_1$ :** premier décile : valeur de la variable telle que 10 % des observations lui soient inférieures.
$F(D_1) = 0,1$ ;
**$D_5$ :** cinquième décile : valeur de la variable telle que 50 % des observations lui soient inférieures.
$F(D_5) = F(Me) = 0,5$
**$D_9$ :** neuvième décile : valeur de la variable telle que 90 % des observations lui soient inférieures.
$F(D_9) = 0,9$ ;

**3)** Les centiles :

en nombre de 99 notés $C_1$, $C_2$,.........$C_{99}$, ils partagent la population en 100 groupes d'effectif égal. Ils sont aussi appelés les percentiles.

**a) Exemple :**

**$C_1$** : premier centile : valeur de la variable telle que 1 % des observations lui soient inférieures.
$F(C_1) = 0,01$ ;
**$C_{50}$** : cinquantième décile : valeur de la variable telle que 50 % des observations lui soient inférieures.
$F(C_{50}) = F(Me) = 0,5$
**$C_{99}$** : quatre vingt dix-neuvième décile : valeur de la variable telle que 99 % des observations lui soient inférieures. $F(C_{99}) = 0,99$ ;

**4)** Application : Reprenant l'exemple précédent :

| Trajet | effectifs | fréquences | F(X) |
|---|---|---|---|
| [10, 20[ | 13 | 0,13 | 0,13 |
| [20, 40[ | 28 | 0,28 | 0,41 |
| [40, 50[ | 21 | 0,21 | 0,62 |
| [50, 80[ | 24 | 0,24 | 0,86 |
| [80, 100[ | 14 | 0,14 | 1 |
| **total** | **100** | **1** | **---** |

$F(10) = 0$ ; $F(20) = 0{,}13$ ; $F(40) = 0{,}41$ ; $F(50) = 0{,}62$ ; $F(80) = 0{,}86$ $et$ $F(100) = 1$.
Cherchons le premier quartile, le sixième décile et le deuxième centile.
- $F(Q_1)=0{,}25$

F (……) = ………... < 0,25

F (…....) = ………… > 0,25

…… ---------------- ………

$Q_1$ **---------------- 0,25**

…… ----------------- ………

À l'aide d'une interpolation linéaire, on aura $Q_1$ =28,571 Km/j.

- $F(D_6)=0{,}6$

F (…..) = …………. < 0,6

F (…..) = ………… > 0,6

……. ---------------- ……..

$D_6$ **---------------- 0,6**

……. ----------------- …….

À l'aide d'une interpolation linéaire, on aura $D_6$ =60 Km/j.

- $F(C_2)=0{,}02$

F (……) =………. < 0, 02

F (……) = ………... > 0, 02

…… ---------------- ……..

$C_2$ **----------------- 0,02**

…… ----------------- ……..

À l'aide d'une interpolation linéaire, on aura $C_2$ =11,538 Km/j.

## Section 3 : MESURE DE LA DISPERSION AUTOUR DE LA MOYENNE

**1)** variance et écart type : on appelle variance d'une variable, la moyenne arithmétique des carrés des écarts entre les valeurs de cette variable et leur moyenne.

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

On peut écrire la variance sous une autre forme appelée formule développée :

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

a) **Démonstration :**

$$V(X) = \sum_{i=1}^{N} f_i(x_i - \bar{X})^2 = \sum_{i=1}^{N}[(f_i x_i)^2 - 2\bar{X}(f_i x_i) + f_i(\bar{X})^2] = \sum_{i=1}^{N} f_i x_i^2 - 2\bar{X}\sum_{i=1}^{N}(f_i x_i) + (\bar{X})^2\sum_{i=1}^{N} f_i$$

$$= \sum_{i=1}^{N}(f_i x_i^2) - 2(\bar{X})^2 + (\bar{X})^2 = \sum_{i=1}^{N}(f_i x_i^2) - (\bar{X})^2$$

b) **Remarque :**

- Pour le cas continu, $x_i$ sera remplacé par $c_i$.
- Cette formule développée de la variance est plus facile à retenir et aussi plus rapide à calculer.
- Si $X = b$ avec $b$ est une constante alors $V(X) = V(b) = 0$
- Si $Y = aX + b$ alors $V(Y) = V(aX + b) = V(aX) + V(b) = V(aX) = \sum_{i=1}^{N} f_i(ax_i)^2 - (a\bar{X})^2$

$V(aX + b) = a^2 \sum_{i=1}^{N} f_i(ax_i)^2 - (a\bar{X})^2 = a^2 [\sum_{i=1}^{N}(f_i x_i^2) - (\bar{X})^2] = a^2 V(X)$ .

c) **L'écart type :**

La variance est exprimée dans le carré de l'unité de la variable, C'est pour cela elle n'est pas facile à interpréter ce qui nous amène à calculer sa racine carrée appelé : **écart type** et noté :

$$\sigma_X = \sqrt{V(X)}$$

L'écart type s'exprime dans la même unité que la variable. Plus l'écart type est grand, plus la dispersion des observations autour de la moyenne est forte.

**2)** Le coefficient de variation :

L'écart type et les indicateurs de tendance centrale sont exprimés dans la même unité de mesure de la variable, ainsi pour comparer la dispersion de deux ou plusieurs distributions exprimées dans des unités différentes, il est indispensable d'utiliser un indicateur de dispersion indépendant de l'unité de mesure, on utilise le coefficient de variation noté : $CV(X)$, c'est un nombre sans unité définit par : $CV(X) = \frac{\sigma_X}{\bar{X}}$.

a) **Application :** Reprenant l'exemple des représentants :

| **Trajet** | **Effectifs $n_i$** | **Centre des classes $c_i$** | $n_i(c_i)^2$ | **Fréquences $f_i$** | $c_i - \bar{X}$ | $(c_i - \bar{X})^2$ | $f_i(c_i - \bar{X})^2$ |
|---|---|---|---|---|---|---|---|
| [10, 20[ | 13 | 15 | ……. | 0,13 | ……. | ……. | ……. |
| [20, 40[ | 28 | 30 | ……. | 0,28 | ……. | ……. | ……. |
| [40, 50[ | 21 | 45 | ……. | 0,21 | ……. | ……. | ……. |
| [50, 80[ | 24 | 65 | ……. | 0,24 | ……. | ……. | ……. |
| [80, 100[ | 14 | 90 | ……. | 0,14 | ……. | ……. | ……. |
| **total** | **100** | **---** | **285450** | **1** | **---** | **--** | **550,5** |

$$V(X) = \frac{1}{N}\sum_{i=1}^{N} n_i c_i^2 - (\bar{X})^2 = \sum_{i=1}^{N} f_i(c_i - \bar{X})^2 = \frac{\dots}{\dots} - (\ \dots\ )^2 = 550{,}50$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{\ \dots\ } = 23{,}463\ km/j$$

$$CV(X) = \frac{\sigma_X}{\bar{X}} = \frac{\dots}{\dots} = 0{,}488\ km/j$$

## Section 4 : MOMENTS D'UNE SERIE STATISTIQUE

**1)** Les moments non centrés :

$$m_r(X) = \frac{1}{N}\sum_{i=1}^{N} n_i x_i^r = \sum_{i=1}^{N} f_i x_i^r$$

- Le moment non centré d'ordre 1 est égale à $m_1(X) = \overline{X}$
- Le moment non centré d'ordre 2 est égale à $m_2(X) = \frac{1}{N}\sum_{i=1}^{N} n_i x_i^2 = \overline{X^2}$

**2)** Les moments centrés :

$$\mu_r = \frac{1}{N}\sum_{i=1}^{N} n_i (x_i - \bar{X})^r = \sum_{i=1}^{N} f_i (x_i - \bar{X})^r$$

- Le moment centré d'ordre 1 est égale à $\mu_1 = \frac{1}{N}\sum_{i=1}^{N} n_i (x_i - \bar{X})^1 = \sum_{i=1}^{N} f_i (x_i - \bar{X})^1 = 0$

- Le moment centré d'ordre 2 est égale à $\mu_2 = V(X) = \frac{1}{N}\sum_{i=1}^{N} n_i (x_i - \bar{X})^2 = \sum_{i=1}^{N} f_i (x_i - \bar{X})^2$

## Section 6 : INDICATEURS DE FORME

Pour avoir une idée satisfaisante et plus précise sur la forme de la distribution, il est nécessaire de calculer les indicateurs de forme. On distingue les indicateurs d'asymétrie et les indicateurs d'aplatissement. Ces indicateurs sont des nombres sans unités de mesure.

**1)** L'asymétrie :

Une distribution est dite symétrique si les observations se répartissent dans la même proportion de part et d'autre des trois valeurs centrales (mode, médiane et moyenne). Les mesures d'asymétrie permettent de quantifier le degré de déviation de la forme de la distribution par rapport à une distribution symétrique.

a) **Comparaison de la moyenne, médiane et mode :**

Selon les positions respectives des paramètres, on peut donner une première impression sur l'allure ou la forme de la distribution :

- si $Mo = Me = \bar{X}$ alors la distribution est symétrique ;
- si $Mo < Me < \bar{X}$ alors la distribution est asymétrique étalée vers la droite ;
- si $Mo > Me > \bar{X}$ alors la distribution est asymétrique étalée vers la gauche.

b) **Le coefficient de Fisher** : noté $\gamma_1 = \frac{\mu_3}{\sigma_X^3}$

Si : $\gamma_1 < 0$ : distribution étalée vers la gauche ;
Si $\gamma_1 = 0$ : distribution symétrique ;
Si $\gamma_1 > 0$ : distribution étalée vers la droite.

c) **Le coefficient de Pearson** :

Noté $C_p$ et il est basé sur la moyenne, le mode et l'écart type : $C_p = \frac{\bar{X} - M_O}{\sigma_X}$

Si $C_p < 0$ : distribution étalée vers la gauche ;
Si $C_p = 0$ : distribution symétrique ;
Si $C_p > 0$ : distribution étalée vers la droite.

**2)** aplatissement :

Une distribution est d'autant plus « plate » que la dispersion des observations autour des valeurs centrales est forte.

a) le coefficient de Pearson : noté $\beta$ et définit par : $\beta = \frac{\mu_4}{\sigma_X^4}$

Si $\beta < 3$ : distribution hyponormale (plus aplatie que la normale) ;
Si $\beta = 3$ : distribution normale ;
Si $\beta > 3$ : distribution hypernormale (moins aplatie que la normale).

b) le coefficient de Fisher : noté $\gamma_2 = \beta - 3$

Si $\gamma_2 < 0$ : distribution hyponormale (plus aplatie que la normale) ;
Si $\gamma_2 = 0$ : distribution normale ;
Si $\gamma_2 > 0$ : distribution hypernormale (moins aplatie que la normale).

**3)** Application :

Reprenant l'exemple des représentants et étudiant la forme de cette distribution.
1/ la symétrie :
- On sait que Mo = 43,50 Km/j; Me = 44,286 Km/j et $\bar{X} = 48\ km/j$
En les comparant on a : Mo < Me < $\bar{X}$ alors la distribution est asymétrique étalée vers la droite.

| Trajet | Effectifs $n_i$ | Centre des classes $c_i$ | $(c_i - \bar{x})$ | $(ci - \bar{x})^3$ | $n_i(ci - \bar{x})^3$ | $n_i(ci - \bar{x})^4$ |
|---|---|---|---|---|---|---|
| [10, 20[ | 13 | 15 | ……. | ……. | ……. | ……. |
| [20, 40[ | 28 | 30 | ……. | ……. | ……. | ……. |
| [40, 50[ | 21 | 45 | ……. | ……. | ……. | ……. |
| [50, 80[ | 24 | 65 | ……. | ……. | ……. | ……. |
| [80, 100[ | 14 | 90 | ……. | ……. | ……. | ……. |
| **total** | **100** | **---** | --- | --- | **524100** | **63926250** |

$$\mu_3 = \frac{1}{N}\sum_{i=1}^{5} n_i(x_i - \bar{X})^3 = \sum_{i=1}^{5} f_i(x_i - \bar{X})^3 = \frac{\dots}{\dots} = 5241$$

- le coefficient de Fisher : $\gamma_1 = \frac{\mu_3}{\sigma_X^3} = \frac{\dots}{\dots} = 0{,}406$

On constate que $\gamma_1 > 0$ : distribution étalée vers la droite.

- le coefficient de Pearson : $C_p = \frac{\bar{X} - M_O}{\sigma_X} = \frac{\dots}{\dots} = 0{,}192$

Comme $C_p > 0$ : distribution étalée vers la droite.

2/ L'aplatissement *:*

$$\mu_4 = \frac{1}{N}\sum_{i=1}^{5} n_i(x_i - \bar{X})^4 = \sum_{i=1}^{5} f_i(x_i - \bar{X})^4 = \frac{\dots}{\dots} = 639262{,}5$$

Le coefficient de Pearson : noté $\beta$ et définit par : $\beta = \frac{\mu_4}{\sigma_X^4} = \frac{\dots}{\dots} = 2{,}109$

On remarque que $\beta < 3$ : distribution hyponormale (plus aplatie que la normale).
Le coefficient de Fisher : noté ; $\gamma_2 = \beta - 3 = \dots - 3 = -0{,}89$
Puisque $\gamma_2 < 0$ alors la distribution hyponormale (plus aplatie que la normale).

Chapitre 3

# ÉVÉNEMENTS ALEATOIRES ET CONCEPT PROBABILITE

## Section 1 : **Événements aléatoires :**

**1)** Expériences aléatoires :

On dit qu'une expérience E est une expérience aléatoire si seulement si on n'est pas certain du résultat obtenu en d'autre terme il y a plusieurs cas possibles.

- Exemple: lorsqu'on jette un dé, l'ensemble des résultats est {1, 2, 3, 4, 5, 6}.

- Les résultats d'une expérience sont appelés : des éventualités.

**2)** Ensemble fondamental :

On l'appelle aussi espace fondamental ou encore univers des possibles noté : Ω ; c'est l'ensemble de tout les résultats possible associés à une expérience :$\Omega = \{W_i \,;\, i \in I\}$.

**a)** **Remarque :** l'ensemble Ω peut prendre plusieurs formes :

- Ω peut être fini.

- Exemple : Le jet d'une pièce de monnaie : $\Omega = \{pile\,;face\}$

- Ω peut être infini dénombrable.

-Exemple : Le jet d'une pièce de monnaie et on s'arrête lorsqu'on obtient pile au premier jet :
Ω={p, fp, ffp, fffp,…fffffff….p, ………………}

- Ω peut être infini non dénombrable.

- Exemple : Soit une expérience qui consiste à lancer une flèche sur le plan du tableau et on s'intéresse au point de contact entre la flèche et le tableau. Dans ce cas il y a une infinité de point ; Ω= {l'ensemble de tout le plan}

**3)** Evènement aléatoire :

On appelle évènement aléatoire un sous ensemble ou une partie de l'ensemble des résultats possibles dont on peut dire suite à la réalisation de l'expérience si elle est effectuée ou non.

- Exemple : Soit un espace fondamental Ω= {1, 2, 3, 4, 5, 6} ; donner les ensembles associé aux évènements suivant :

1/ Obtenir le nombre 1 ; soit $W_1 = \{ \ldots \}$

2/ Obtenir un nombre paire ; soit $W_2 = \{ \ldots ; \ldots ; \ldots \}$

3/ Obtenir un nombre impaire ; soit $W_3 = \{ \ldots ; \ldots ; \ldots \}$

### a) Définition de quelques évènements :

- Evènement élémentaire : c'est un sous ensemble de Ω qui est composé d'un seul élément (réduit à un singleton).
- Evènement certain : il est confondu avec l'ensemble des possibles Ω.
-Evènement impossible : il est confondu avec le sous ensemble vide $\emptyset$ .

### b) opérations sur les évènements :

- Union : noté par $\bigcup$ ; $A \bigcup B$ expliqué en terme d'évènement ; c'est soit A soit B ; soit les deux simultanément qui se réalise. Elle se traduit par "OU".
- Intersection : noté par $\bigcap$ ; $A \bigcap B$ s'explique en terme d'évènement par la réalisation simultanée des deux évènements A et B. Elle se traduit par "ET".
- Complémentaire : L'évènement complémentaire à l'évènement A est l'évènement $\bar{A}$ qui se traduit par la non réalisation de A.
- Inclusion : noté par $\subset$, si on a deux évènements A et B tel que $A \subset B$ en terme d'évènement la réalisation de A entraîne la réalisation de B.
- Evènements incompatibles : appelée encore distincts ou encore disjoints. On dit que deux évènements A et B sont incompatibles si $A \bigcap B = \emptyset$.

Remarque :

La relation d'incompatibilité est une relation entre évènements.

- Système complet d'évènement : Soit une suite d'évènements $A_1$ , $A_2$, ...,$A_n$ tel que $A_i \cap A_j = \emptyset$ ; $\forall A_i \neq A_j$ $et$ $\bigcup_{i=1}^{n} A_i = \Omega$ ; on dit que la suite $A_i$ constitue une partition de Ω où un système complet d'évènement.

## Section 2 : **Concept de probabilité :**

**1)** Espace probabilisable :

Soit l'espace probabilisable $(\Omega; \varphi)$ où φ est l'ensemble des évènements qui appartiennent à Ω et W un évènement. On appelle probabilité sur $(\Omega; \varphi)$ ; l'application P définie par : $P: \varphi \to [0;1]$

$$W \to P(W)$$

**OBJECTIF :** associer à chaque évènement w de l'ensemble des possibles Ω un réel appartenant à $[0;1]$ qu'on appellera probabilité de w et ayant les propriétés suivantes : **Axiomes de probabilités** :

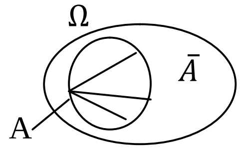
- $P(\Omega) = 1$
- $P(w_1 \cup w_2) = P(w_1) + P(w_2)$ si $w_1 \cap w_2 = \emptyset$
- $0 \leq P(w)$ ; $\forall\ w \subset \Omega$

**2)** Notions utiles :

- ❖ la probabilité de l'évènement certain est égale à : $P(\Omega) = 1$
- ❖ la probabilité de l'évènement impossible est égale à : $P(\emptyset) = 0$
- ❖ la probabilité de l'évènement complémentaire est égale à : $P(\bar{A}) = 1 - P(A)$.

**Démonstration :** $A \cap \bar{A} = \emptyset \; et \; A \cup \bar{A} = \Omega$ donc

$P(A \cup \bar{A}) = P(\Omega) = 1 = P(A) + P(\bar{A}) = 1$



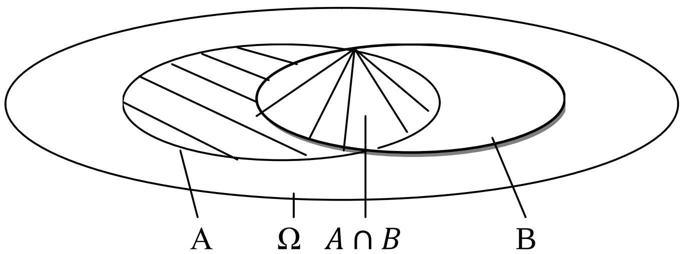

**3)** Remarque :

- ❖ Si un évènement en implique un autre, sa probabilité est plus petite :

$$A \subset B \implies P(A) \leq P(B)$$

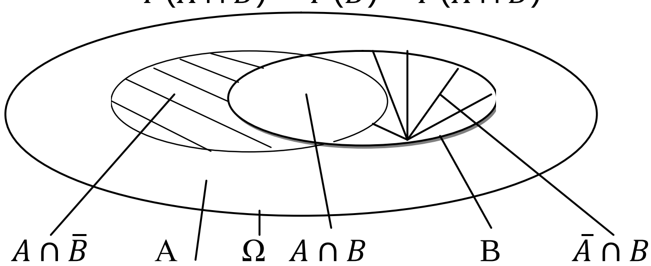
- ❖ La probabilité de l'union de 2 évènements s'obtient par la formule de Poincaré :

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ car $(A \cap B) \subset A \;\; et \;\; (A \cap B) \subset B$




- ❖ Soit Ω un espace fondamental et soient A et B deux évènements distincts ou non ; on a :

$$P(\bar{A} \cap B) = P(B) - P(A \cap B)$$




**Démonstration :** $P(B) = P(\Omega \cap B) = P([A \cup \bar{A}] \cap B) = P(A \cap B) + P(\bar{A} \cap B)$

Puisque $B \subset \Omega$ et $A \; et \; \bar{A}$ forment un système complet d'évènement

- ❖ Deux évènements A et B ayant la même probabilité : $P(A) = P(B)$ sont équiprobables.

**4)** Formule de la place :

Lorsqu'on est en présence d'évènements élémentaires, équiprobables et distincts ; la probabilité d'un évènement est donnée par la formule suivante :

$$P(A) = \frac{Nombre \;\; de \; cas \;\; favorable \; à \; A}{Nombre \quad de \quad cas \quad possible}$$

- ❖ Exemple : Lancer un dé et chercher la probabilité d'avoir comme résultat 1 : $P(1) = \frac{\dots}{\dots}$.

**5)** Loi de probabilité :

Soit un univers des possibles Ω et $\varphi$ l'ensemble des évènements appartenant à Ω.

Une probabilité ou loi de probabilité P est une application de $P: \varphi \to [0;1]$ vérifiant :

**Les axiomatiques de Kolmogorov** :

1/ $P(W) \geq 0$

2/ $P(\Omega) = 1$

3/ si $w_1, w_2 \dots w_n$ est une suite d'évènement mutuellement exclusifs alors : $P(\bigcup_{i=1}^{n} w_i) = \sum_{i=1}^{n} P(w_i)$

- Toute fonction qui vérifie ces trois propriétés est une loi de probabilité.

**6)** Probabilité conditionnelle :

a) **Définition :**

On définit la probabilité conditionnelle de l'évènement B par rapport à l'évènement A (ou probabilité de B sachant A) par :

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \quad avec\ P(B) \neq 0 \ ; \ P(B/A) = \frac{P(B \cap A)}{P(A)} \quad avec\ P(A) \neq 0$$

❖ Cette définition permet d'obtenir :

$$P(A \cap B) = P(A/B) \times P(B) = P(B/A) \times P(A) \quad avec\ P(A) \neq 0\ \ et\ P(B) \neq 0$$

✓ Exemple : Considérons l'expérience qui consiste à jeter un dé. Si avant l'expérience on vous donne l'information supplémentaire suivante : le résultat est un nombre impair. Par suite, connaissons cette information qu'elle est la probabilité d'avoir un ?

Soit l'évènement A : " avoir un" et l'évènement B : "avoir un nombre impair".

✓ Réponses : Comme …. ⊂ ….. donc $P(\dots . \cap \dots .) = P(\dots .)$

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\dots)}{P(\dots)} avec\ P(B) \neq 0$$

$$P(A) = \frac{\dots\dots}{\dots\dots} \ ; \ P(B) = \frac{\dots\dots}{\dots\dots} \ \text{ et } \ P(A/B) = \frac{P(A)}{P(B)} = \frac{\frac{\dots}{\dots}}{\frac{\dots}{\dots}} = \frac{\dots\dots}{\dots}$$

b) **Indépendance stochastique:**

Deux évènements sont stochastiquements indépendants, si la donnée de l'information sur la réalisation de B (respectivement de A) n'agit en rien sur la probabilité de réalisation de A (respectivement de B). C'est-à-dire : $P(A) = P(A/B) = \frac{P(A \cap B)}{P(B)} \ \ avec\ P(B) \neq 0$

Donc, on déduit que si les évènements A et B sont indépendants alors :

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

On dit que A et B sont stochastiquements indépendants.

Remarque : Si A et B sont stochastiquement indépendants alors $\bar{A}$ et $\bar{B}$ ; A et $\bar{B}$ ; $\bar{A}$ et B sont indépendants.

c) **Les théorèmes de Bayes :**

❖ ***Théorème de Bayes*** : $P\left({}^{A}/_{B}\right) = \frac{P(A\cap B)}{P(B)} = \frac{P\left({}^{B}/_{A}\right)\times P(A)}{P(B)}$ $avec\ P(B) \neq 0$

❖ ***Formule de Bayes Généralisée*** :

On considère un système complet d'évènements : d'évènements $A_1$ , $A_2$, ...,$A_n$ tel que $A_i \cap A_j = \emptyset$ $\forall A_i \neq A_j\ et \bigcup_{i=1}^{n} A_i = \Omega$ et B un évènement de $\varphi$ ; par définition on a :

$$P\left({}^{A_i}/_{B}\right) = \frac{P(A_i\cap B)}{P(B)} = \frac{P\left({}^{B}/_{A_i}\right)\times P(A_i)}{P(B)} \text{ avec } P(B) \neq 0$$

Ou encore $P\left({}^{A_i}/_{B}\right) = \frac{P(A_i\cap B)}{P(B)} = \frac{P\left({}^{B}/_{A_i}\right)\times P(A_i)}{\sum_{j=1}^{n} P\left({}^{B}/_{A_j}\right)\times P(A_j)}$ $telque\ j \in \{1;2 \dots n\}$

$$P(B) = P(B\cap \Omega) \text{ et on sait que} \Omega = \bigcup_{j=1}^{n} A_j\ donc\ P(B) = P\left(\bigcup_{j=1}^{n} [B\cap A_j]\right) = \sum_{j=1}^{n} P\left({}^{B}/_{A_j}\right)\times P(A_j)$$

❖ <u>Exemple :</u>

12 appareils sont en exploitations, 3 d'entre eux sont fabriquées par l'usine I, 4 par l'usine II et 5 par l'usine III.

Les appareils provenant de l'usine I passent à l'essai avec une probabilité de 0,9.

Les appareils provenant de l'usine II passent à l'essai avec une probabilité de 0,8.

Les appareils provenant de l'usine III passent à l'essai avec une probabilité de 0,75.

Déterminer la probabilité pour qu'un appareil choisi au hasard passe à l'essai.

❖ <u>Réponse :</u>

Soit l'évènement B : l'appareil choisi passe l'essai.

Soit les évènements $A_j$ : l'appareil provient de l'usine j. (j=1 ;2 ;3).

On sait que : $P(\cdots/\dots) = \cdots$ ; $P(\cdots/\dots) = \cdots$ $et\ P(\cdots/\dots) = \cdots$

On a aussi : $P(\dots) = \frac{\dots}{\dots}$ ; $P(\dots) = \frac{\dots}{\dots}$ $et\ P(\dots) = \frac{\dots}{\dots}$

Donc : $P(B) = \sum_{i=1}^{3} P\left({}^{B}/_{A_i}\right)\times P(A_i)$

$$= P\left({}^{B}/_{A_1}\right)\times P(A_1) + P\left({}^{B}/_{A_2}\right)\times P(A_2) + P\left({}^{B}/_{A_3}\right)\times P(A_3)$$

$$= \cdots \times \frac{\dots}{\dots} + \dots \times \frac{\dots}{\dots} + \dots \times \frac{\dots}{\dots} = \cdots + \cdots + \cdots = 0{,}8041$$

Chapitre 4

# LES VARIABLES ALEATOIRES REELLES

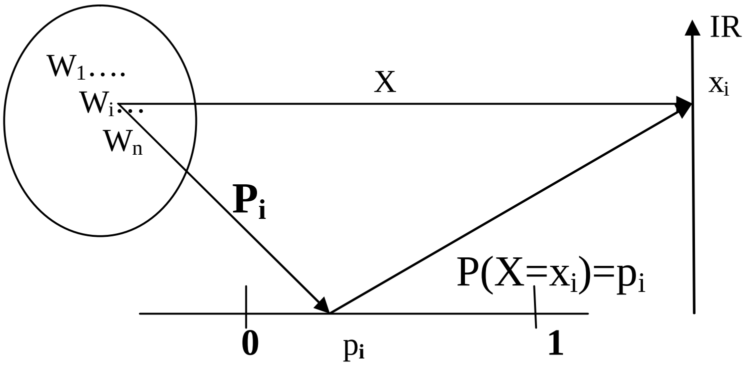
## I- Introductions :

Une variable dont la valeur est déterminer en fonction du résultat d'une expérience aléatoire est appelée variable aléatoire (V.A).

Une variable aléatoire (VA) est donc une fonction à valeurs défini sur l'ensemble fondamental.

Autrement dit, une variable aléatoire (V.A) réelle X est une application de $\Omega$ dans IR ; $\begin{array}{l} X:\Omega \rightarrow IR \\ W_i \rightarrow X_i \end{array}$

On distingue généralement les variables aléatoires dites discrètes et celle qualifié de continue.




## II- Variables aléatoires discrètes :

### Section 1 : Définition :

Une variable aléatoire est dite discrète si l'ensemble de ces valeurs $X(\Omega)$ est dénombrable (fini ou infini) ; comme par exemple : pour tout X, nous avons une application de $\Omega$ dans IN : $\Omega \rightarrow IN$ .

**1)** Exemple **:** si on lance une pièce de monnaie deux fois $\Omega$ est égale à : $\Omega = \{PP, PF, FP, FF\}$ à partir de cet ensemble on peut définir diverses variables aléatoires dont par exemple :

**X** : Nombre total de pile

**Y** : Avoir un si pile apparait en premier lieu sinon zéro.

| Eventualité | X | Y |
|---|---|---|
| PP | 2 | 1 |
| PF | 1 | 1 |
| FP | 1 | 0 |
| FF | 0 | 0 |

Valeurs des V.As

On remarque X est une V.A prenant des valeurs dans l'ensemble $\{0;1;2\}$ alors que Y est aussi une VA définie sur le même ensemble fondamental que X mais prend ces valeurs dans l'ensemble $\{0;1\}$.

Section 2 : **Loi de probabilité :**

La loi de probabilité d'une variable X, on dit aussi la distribution de X, est le mode de répartition de la probabilité de $\Omega$ sur l'ensemble des observations.

Il en résulte que connaitre la loi de probabilité de X ; c'est être capable d'attribuer une probabilité à chaque valeur $x_i$ de la V.A.

**Remarque :** une loi de probabilité n'est établie que si $\sum_i p_i = 1$

La loi de probabilité relative à la V.A.D X est comme suit ;

**Distribution de X ou Loi de probabilité de X**

| $x_i$ | 0 | 1 | 2 | total |
|---|---|---|---|---|
| $P(x_i)$ | $\frac{\dots}{\dots}$ | $\frac{\dots}{\dots}$ | $\frac{\dots}{\dots}$ | 1 |

Section 3 : **Fonction de répartition :**

La fonction de répartition d'une variable aléatoire discrète (V.A.D) X est la fonction F donnée pour tout x de l'ensemble de définition de X par : $F(x) = P(X \leq x)$ ; en d'autre terme c'est l'application suivante :

$$F : IR \to [0;1]$$
$$x \mapsto F(x) = P(X \leq x)$$

càd

$$F(x) = \begin{cases} 0 & si\ x < x_1 \\ P(X \leq x_1) & si\ x_1 \leq x < x_2 \\ \vdots & \\ P(X \leq x_k) & si\ x_{k-1} \leq x < x_k \end{cases}$$

- La fonction de répartition F vérifie les trois propriétés suivantes :

  1/ F est une fonction croissante.

  2/ $\lim_{x \to -\infty} F(x) = 0$ *et* $\lim_{x \to +\infty} F(x) = 1$

  3/pour une V.A.D X, F est une fonction en escalier continue à droite.

**1)** Remarque :

Lorsque l'ensemble de définition de X est $E = \{x_1, x_2, x_3, \dots, x_n\}$ on a les résultats suivants :

1/ F(x)=0 pour tout $x \in ]-\infty; x_1[$

2/ F garde la même valeur $F(x_i)$ sur tout l'intervalle $[x_i; x_{i+1}[$

3/ Au point $x_i$ ; F fait un saut en hauteur égale à la probabilité $P(X=x_i)$

4/ On a en particulier $P(X = x_i) = F(x_i) - F(x_{i-1})$

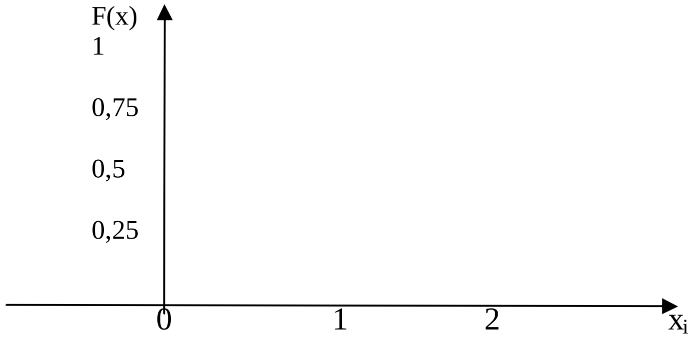
**2)** Exemple : Reprenant l'exemple précédent :

F(0)=F(X=0)=$P(X \leq 0) = P(X = 0) = \frac{\cdots}{\cdots}$

$$F(x) = \begin{cases} \cdots & si\ x < 0 \\ \cdots & si\ 0 \leq x < 1 \\ \cdots & si\ 1 \leq x < 2 \\ \cdots & si\ x \geq 2 \end{cases}$$

F(1)=F(X=1)=$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{\cdots}{\cdots} + \frac{\cdots}{\cdots} = \frac{\cdots}{\cdots}$

F(2)=F(X=2)=$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{\cdots}{\cdots} + \frac{\cdots}{\cdots} + \frac{\cdots}{\cdots} = 1$




Section 4 : **Espérance mathématique :**

**1)** Définition :

L'espérance mathématique E(X) de la V.A X est la moyenne arithmétique des valeurs possibles pondérées par les probabilités correspondantes. Dans le cas d'une VA discrètes on a :

$$E(X) = \sum_{i=1}^{n} x_i\, P(X = x_i)$$

Il est facile de monter que l'espérance mathématique d'une V.A (aussi bien continue que discrète) vérifie les propriétés suivantes :

1/ Si b est une constante alors $E(b) = b$

2/ Si a et b deux constantes et X une VA alors $E(a\,X + b) = a\,E(X) + b$

3/ soient deux VAs $X$ et $Y$ ; on a $E(\,X \pm Y) = E(X) \pm E(Y)$

Exemple : suite du même exemple :

$$E(X) = \sum_{i=1}^{3} x_i\, P(X = x_i) = \cdots \times \frac{\cdots}{\cdots} + \cdots \times \frac{\cdots}{\cdots} + \cdots \times \frac{\cdots}{\cdots} = 1$$

Section 5 : **Variance et écart type :**

**1)** La variance : La variance de la VA X est : V(X) mesure l'amplitude de la variabilité de X, elle est définie comme l'espérance des carrés des écarts à l'espérance :

$$V(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2$$

Pour les VA discrètes on a plus précisément :

$$V(X) = \sum_{i=1}^{n}[x_i - E(X)]^2\, P(X = x_i) = \sum_{i=1}^{n} x_i^2\, P(X = x_i) - \left[\sum_{i=1}^{n} x_i\, P(X = x_i)\right]^2$$

**2)** L'écart type :

L'écart type de la V.A X ( $\sigma_X$ ) est la racine carrée de la variance :

$$\sigma_X = \sqrt{V(X)} = \sqrt{E(X^2) - [E(X)]^2}$$

a) **Remarque :** on montre que la variance d'une V.A (aussi bien continue que discrète) vérifie les propriétés suivantes :

1/ Si b est une constante alors $V(b) = 0$

2/ Si a et b deux constantes et X une VA alors $V(a\,X + b) = a^2V(X)$

b) **Exemple :** suite du même exemple :

$$V(X) = (\ldots - \cdots)^2\left(\frac{\cdots}{\cdots}\right) + (\ldots - \cdots)^2\left(\frac{\cdots}{\cdots}\right) + (\ldots - \cdots)^2\left(\frac{\cdots}{\cdots}\right) = \frac{\cdots}{\cdots}$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{\frac{\cdots}{\cdots}} = \frac{\cdots}{\sqrt{\ldots}}$$

## Section 6 : **Les moments centrés et non centrés :**

On appelle moment non centré d'ordre $K$ $(K \in IN)$ de X, et noté : $E(X^K)$, le nombre réel, s'il existe, définie par : $m_K = \sum_i [x_i^K P(x_i)] = \sum_i x_i^K P_i = E(X^K)$

- Cas particulier : $m_0 = \sum_i P_i = E(X^0) = 1$ ; $m_1 = \sum_i x_i P_i = E(X)$; $m_2 = \sum_i x_i^2 P_i = E(X^2)$

Les moments centrés d'ordre K, noté $\mu_r(X)$, l'expression suivante :

$$\mu_K(X) = E[X - E(X)]^K = \sum_i [x_i - E(X)]^K\, P(X = x_i)$$

- Cas particulier $\mu_1(X) = E[X - E(X)]^1 = E(X) - E[E(X)] = E(X) - E(X) = 0$ :

$$\mu_2 = E[X - E(X)\ ]^2 = E[X^2 - 2XE(X) + E(X)^2] = E(X^2) - 2[E(X)]^2 + [E(X)]^2$$

$$\mu_2(X) = E[X - E(X)]^2 = E(X^2) - [E(X)]^2 = V(X)$$

## Section 7 : **Fonction génératrice des moments:**

Il est parfois difficile de déterminer la distribution d'une probabilité d'une V.A par les méthodes directes. Néanmoins en faisant correspondre à chaque variable une fonction dont la connaissance est équivalente à la connaissance de la distribution de probabilité, nous pourrons identifier la loi.

**1)** Définition :

On appelle fonction génératrice des moments d'une V.A X noté par $M_X(t)$, l'expression suivante : $M_X(t) = E(e^{xt}) = \sum_i e^{tx_i} P(X = x_i)$

Cette fonction nous permet de retrouver facilement les moments d'une loi, en effet, le n$^{\text{ème}}$ moment est égale au n$^{\text{ème}}$ dérivée de $M_X(t)$ évalué au point t=0 :

$$M_X^{(1)} = \frac{\partial M_X(t)}{\partial t} = \frac{\partial(e^{tx_i}P_i)}{\partial t} = \sum_i \left(\frac{\partial e^{tx_i}}{\partial t}\right) P_i = \sum_i x_i e^{tx_i} P_i$$

Au point t=0 on aura : $M_X^{(1)} = \frac{\partial M_X(t)}{\partial t}|t_{=0} = \sum_i x_i P_i = m_1 = E(X)$

$$M_X^{(2)} = \frac{\partial^2 M_X(t)}{\partial t^2} = \frac{\partial M_X^{(1)}(t)}{\partial t} = \sum_i \left(\frac{\partial e^{tx_i}}{\partial t}\right) x_i P_i = \sum_i x_i^2 e^{tx_i} P_i$$

Au point t=0 on a : $M_X^{(2)} = \frac{\partial^2 M_X(t)}{\partial t^2}|t_{=0} = \sum_i x_i^2 P_i = m_2 = E(X^2)$

.
.
.

$$M_X^{(K)} = \frac{\partial^K M_X(t)}{\partial t^K} = \frac{\partial M_X^{(K-1)}(t)}{\partial t} = \sum_i \left(\frac{\partial e^{tx_i}}{\partial t}\right) x_i^{K-1} P_i = \sum_i x_i^{K-1} P_i$$

Au point t=0 on a : $M_X^{(K)} = \frac{\partial^K M_X(t)}{\partial t^K}|t_{=0} = \sum_i x_i^K P_i = m_K = E(X^K)$

**2)** Application :

A la caisse d'un super marché, le client peut payer en espèces ou par carte bancaire. Pour chaque client se présentant à la caisse, on définit X telle que : $X = 1$ Si le mode de payement est par carte bancaire et $X = 0$ sinon.

La loi de probabilité de X est définie par la fonction de probabilité suivante :

$$\begin{cases} P(X = x) = p^x(1-p)^{1-x} & avec\ x \in \{0;1\} \\ 0 & sinon \end{cases}$$

1/ Etablir le moment non centré x de X tel que $x \in IN^*$, en déduire $V(X)$

2/ Calculer la fonction génératrice des moments d'une V.A $X$ puis déduire $V(X)$ .

**3)** Réponse :

# III- Variables aléatoires continues :

## Section 1 : **Définition :**

Une V.A est dite continue si l'ensemble de ces valeurs X(Ω) est non dénombrable en d'autre terme elle peut prendre ces valeurs dans un intervalle. C'est une application de Ω dans IR ; Ω →IR.

Beaucoup de mesures de quantités s'exprime en terme de V.A.C comme : la durée d'un appelle téléphonique ; le poids d'un individu ou encore la direction du vent qui peut prendre n'importe quel angle entre 0° et360°…………

## Section 2 : **Fonction de répartition :**

On appelle fonction de répartition d'une V.A.C la fonction F définie par :

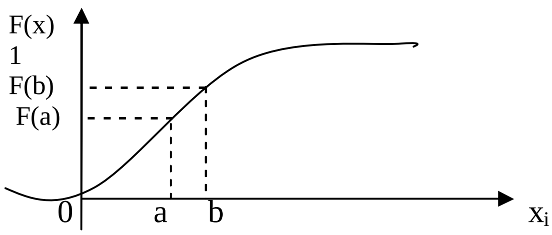
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

Pour un nombre réel x ; la fonction de répartition de X correspond donc à la probabilité que X soit inférieur ou égale à x.

**Remarque :** Lorsque la V.A.C $X$, la probabilité de $X$ soit **exactement** égale à une valeur précise $x$ est nulle : $P(X = x) = 0$.

**1)** Propriétés :

- $0 \leq F(x) \leq 1$ ; $\lim_{x\to-\infty} F(x) = 0$ $et$ $\lim_{x\to+\infty} F(x) = 1$ .
- F est une fonction croissante, continue et dérivable sur $D_f$.
- La représentation graphique d'une fonction de répartition d'une V.A.C se présente comme suit :




La probabilité que la V.A.C X prend une valeur dans l'intervalle $]a; b]$ est :

$$P(X \in ]a; b]) = P(a \leq X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

a- **Exemple :** soit la V.A.C X donné par la fonction de répartition :

$F(x) = \begin{cases} 0 \; si \; x < 2 \\ (x-2)^2 \;\; si \; 2 \leq x \leq 3 \\ 1 \;\; si \; x > 3 \end{cases}$ .Calculer la probabilité que X tombe dans l'intervalle$[2{,}5 \leq X \leq 3{,}5]$.

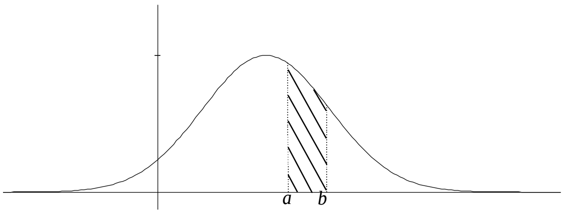
$P(2{,}5 \leq X \leq 3{,}5) = P(X \leq \cdots) - P(X \leq \cdots) = F(\dots) - F(\dots) = \cdots - (\dots - \cdots)^2 = \cdots - \cdots = \cdots$

Section 3 : **Fonction de densité :**

Une fonction est appelée densité de probabilité de la V.A.C X si elle correspond à la dérivée de la fonction de répartition F(x) au point x : $f(x) = \frac{\partial F(x)}{\partial x} = F'(x)$

La probabilité que X prend une valeur comprise entre deux borne a et b est donc égale à :

$$P(a \leq X \leq b) = \int_{-\infty}^{a} f(t)dt - \int_{-\infty}^{b} f(t)dt = \int_{a}^{b} f(t)dt = F(b) - F(a)$$




La surface hachurer représente la probabilité entre a et b càd : $P(a \leq X \leq b)$.

**1)** Remarque :

1/ La loi de probabilité d'une V.A .C est entièrement définie par sa densité de probabilité.

2/ Toute densité de probabilité f vérifie les deux propriétés suivantes :

- $f(x) \geq 0$ pour tout $x$ appartenant au domaine de f $(D_f)$.
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

**2)** Exemple : On continue avec l'exemple précédent : Chercher la fonction de densité

On sait que $(x) = \frac{\partial F(x)}{\partial x} = F'(x)$ , on a alors :

Section 4 : **Espérance mathématique :**

Soit la V.A.C X, prenant ces valeurs sur un intervalle D ; on appelle valeur moyenne ou espérance mathématique de X ; si elle existe ; le nombre : $E(X) = \int_D xf(x)\partial s = \mu$

**Remarque :** dans le cas ou $E(X) = \pm\infty$ ou elle est en fonction de $x$ ; on dit que espérance n'existe pas.

**1)** Propriétés : soit la V.A X et deux constantes a et b.

- $E(b) = b; E(a\,X) = a\,E(X)$ et aussi $E(a\,X + b) = a\,E(X) + b$

**2)** Exemple : suite de la même application : Calculer l'espérance mathématique.

## Section 5 : **Variance et écart type :**

La variance mesure la déviation par rapport à la moyenne$\mu$ :

V(X)=$\sigma_X^2 = \int_D^{.} (x-\mu)^2 f(x)\partial x = \int_D^{.} x^2 f(x)\partial x - \mu^2 = \mathrm{E}(\mathrm{X}^2) - (\mathrm{E}(\mathrm{X}))^2$

L'écart type est la racine carrée de la variance : $\sigma_X = \sqrt{V(X)}$.

**1)** Propriétés : soit X une V.A X et deux constantes a et b.

- $V(a\,X) = a^2\,V(X)$
- $V(aX+b) = E([\mathrm{a\,X+b}] - \mathrm{E}[\mathrm{aX+b}]\,)^2 = E(aX+b-aE(X)-b) = a^2E(X-E(X))^2 = a^2V(X)$

**2)** Application : suite du même exemple : Calculer la variance et l'écart type.

## Section 6 : **Les moments centrés et non centrés :**

On appelle moment non centré d'ordre K (K∈ IN) de X ; le moment rapportés à l'origine, et noté : $E(X^K)$, le nombre réel, s'il existe, définie par : $m_k(X) = E(X^K) = \int_{-\infty}^{+\infty} x^K f(x)dx$.

Les moments centrés d'ordre K (K∈ IN) de X ; est le moment rapportés à la moyenne, noté $\mu_r(X)$ :

$$\mu_K(X) = E[X-E(X)]^K = \int_{-\infty}^{+\infty} (x-\mu)^K f(x)\partial x$$

- Cas particulier $\mu_1(X) = E[X-\mu]^1 = \mu - E[\mu] = \mu - \mu = 0$ :

$$\mu_2(X) = E[X-\mu]^2 = E(X^2) - [\mu]^2 = m_2(X) - m_1(X)^2 = V(X)$$

1) Application : suite du même exemple :

Calculer $m_k(X)$ puis déduire $V(X)$.

## Section 7 : **Fonction génératrice des moments:**

**1)** Définition :

On appelle fonction génératrice des moments d'une V.A X noté par $M_X(t)$, l'expression suivante :

$M_X(t) = E(e^{xt}) = \int_{-\infty}^{+\infty} e^{tx} f(x) \partial x$

Cette fonction nous permet de retrouver facilement les moments d'une loi, en effet, le n[ème] moment est égale au n[ème] dérivée de $M_X(t)$ évalué au point $t = 0$ : $E(X^n) = M_X^{(n)}(0) = \frac{\partial^n M_X(t)}{\partial t^n} |t_{=0}$

**a) Démonstration :**

$$M_X(t) = E(e^{xt}) \ \ or \ \ e^{xt} = \sum_{K=0}^{+\infty} \frac{(xt)^K}{K!}$$

$$M_X(t) = E\left(\sum_{K=0}^{+\infty} \frac{(xt)^K}{K!}\right) = E\left(1 + xt + \frac{x^2t^2}{2!} + \frac{x^3t^3}{3!} \dots\right) = 1 + tE(X) + \frac{t^2E(x^2)}{2!} + \frac{t^3E(x^3)}{3!} \dots$$

$$\frac{\partial M_X(t)}{\partial t} = E(X) + tE(X^2) + \frac{t^2}{2} E(X^3) \dots\dots \Longrightarrow \frac{\partial M_X(t)}{\partial t} |t_{=0} = E(X).$$

**2)** Propriétés :

- $M_X(0) = E(e^{x0}) = E(e^0) = E(1) = 1$
- $M_{aX+b}(t) = E\left(e^{(aX+b)t}\right) = E(e^{aXt+bt}) = E(e^{atx})e^{tb} = e^{tb} M_X(at)$

**3)** Théorème :

Pour X et Y aient la même loi, il est nécessaire et suffisant qu'elles admettent la même fonction génératrice des moments : $M_X(t) = M_Y(t) \iff X \ et \ Y$ ont la même loi.

**4)** Application :

Soit la fonction de $f(x) = \begin{cases} e^{-x} & \text{si } x \geq 0 \\ 0 & \text{ailleurs} \end{cases}$

Déterminer $M_X(t)$ puis en déduire $V(X)$et $M_Y(t)$ sachant que $Y = 2X + 1$.

**5)** Réponse :

## Section 8 : **Transformation d'une variable aléatoire :**

**1)** premième méthode :

Soit X est une V.A de densité $f_X$ ; $Y = g(x)$ tel que $Y$ est une fonction monotone. Si $f$ est continue et $g^{-1}$est à dérivée continue ; alors : $f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left|\frac{\partial g^{-1}(y)}{\partial y}\right| & y \in Y \\ 0 \ \ sionon \end{cases}$

**2)** Deuxième méthode :

En se basant sur la définition de la fonction de répartition de la V.A Y et celle de X et sur la fonction $Y = g(x)$ ; on peut retrouver $F_Y(y)$ et en la dérivant on retrouve la fonction de densité des Y.

**a- Exemple :** suite de l'application :

Chercher la densité de Y sachant que : $g(x) = y = 2x + 1$ de deux méthodes.

**b- Réponse :**

Chapitre 5

# LES LOIS DE PROBABILITE

## I- Introductions :

La plupart des phénomènes aléatoires peuvent être décrits par un petit nombre de lois de probabilité. Donner la loi de probabilité qui décrit au mieux le phénomène aléatoire étudié est beaucoup plus informatif que donner uniquement quelques caractéristiques comme la moyenne ou la variance.

Il est important de connaitre les modèles probabilistes les plus susceptibles de convenir à la description du phénomène aléatoire étudié.

Le processus de modélisation passe par les étapes suivantes :

1. L'observation du phénomène donne une distribution expérimentale ou empirique.
2. L'analyse de la distribution empirique (graphique, moyenne, étc…) donne une première idée sur la nature du phénomène étudié.
3. On choisit ensuite parmi les différentes lois de probabilité celle qui semble convenir le mieux.
4. Utiliser le modèle ainsi obtenu pour la prévision et l'optimisation.

## II- Les lois de probabilités discrètes :

### Section 1 : La loi uniforme :

On dit qu'une V.A X suit la loi uniforme discrète, si pour un nombre finie de valeurs réelles $X(\Omega) = \{1, 2, \dots, n\}$ ; sa loi de probabilité est définie par : $P(X = x_i) = \frac{1}{n}$ ; $\forall$ $i \in \{1; 2; \dots; n\}$.

- L'expérience aléatoire consiste ici à choisir au hasard un objet parmi n objet numérotés de 1 à n et équiprobable (le lancer d'un dé par exemple).

**1)** Caractéristiques:

$$E(X) = \sum_{k=1}^{n} k\,P(X = k) = \frac{1}{n}\sum_{k=1}^{n} k = \frac{1}{n}\;\frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$E(X^2) = \sum_{k=1}^{n} k^2\,P(X = k) = \frac{1}{n}\sum_{k=1}^{n} k^2 = \frac{1}{n}\;\frac{n(n+1)(2n+1)}{6} = \frac{(2n+1)(n+1)}{6}$$

$$V(X) = E[(X - E(X))^2] = E(X^2) - [E(X)]^2 = \frac{(2n+1)(n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)(n-1)}{12}$$

## Section 2 : **Loi de probabilité de Bernoulli:**

Cette loi de probabilité illustre toute expérience aléatoire à deux issues effectuées une seule fois (passer l'examen de stat et avoir la moyenne ou non).

Ces deux résultats on peut les codés par 0 et 1 comme suit : $X \begin{cases} 1 & \text{si l'étudiant a eu sa moyenne} \\ 0 & \text{sinon} \end{cases}$

Soit p la probabilité d'avoir la moyenne ; la loi de Bernoulli est alors : $X \rightsquigarrow \beta(1;p)$ si $X(\Omega) = \{0;1\}$

$$P(X=1) = p \text{ et } P(X=0) = 1-p = q$$

**1)** Caractéristiques:

$$m_K(X) = E(X^K) = \sum_{x=0}^{1} x^K P(X=x) = 0^K \times P(X=0) + 1^K \times P(X=1) = p$$

$$E(X) = \sum_{x=0}^{1} x P(X=x) = 0 \times P(X=0) + 1 \times P(X=1) = p$$

On a: $X(\Omega) = \{0;1\} \Rightarrow X^2(\Omega) = \{0^2;1^2\} = \{0;1\} = X(\Omega) \Rightarrow X^2 = X \Rightarrow E(X^2) = E(X) = p$

$$V(X) = E[(X-E(X))^2] = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p) = p.q$$

La forme symétrique de la loi de Bernoulli est : $P(X=x) = p^x(1-p)^{1-x}$ avec $x \in \{0;1\}$

La fonction génératrice des moments de X est définie par : $M_X(t) = E(e^{xt}) = q + e^t p$

$$M_X(t) = \sum_{x=0}^{1} e^{tx} p^x (1-p)^{1-x} = e^0 p^0 (1-p)^{1-0} + e^1 p^1 (1-p)^{1-1} = (1-p) + e^t p = q + e^t p$$

## Section 3 : **La loi Binomiale :**

**1)** Théorème :

Si $X \rightsquigarrow \beta(n;p)$ c'est équivalent de dire que X est la somme de n V.A indépendants et identiquement distribuer (idd) selon la loi Bernoulli.

- Soit une suite $X_i \,/\, i \in \{1;2;..;n\}$ de n variables de Bernoulli indépendants et X la V.A liée au nombre de "réussite" acquis au terme de n épreuves.
- La loi suivie par X vérifie les trois conditions qui caractérisent la loi Binomiale :
  1. Le nombre d'épreuves est fixé et connue à l'avance ; c'est **n**.
  2. A chaque épreuve, il existe une alternative : "**réussite**" ou "**échec**".
  3. La probabilité **p** d'obtention d'une "réussite" dans différents épreuves est constante. C à d les épreuves sont mutuellement indépendantes.

On dit qu'une V.A X suit la loi Binomiale de paramètre n et p ; notée : $X \rightsquigarrow \beta(n;p) \,/\, X(\Omega) = \{1;2;..;n\}$

$$P(X=k) = C_n^k p^k (1-p)^{n-k}$$

**2)** Caractéristiques :

On considère la suite $X_i\ / i\ \in \{1; 2; ..; n\}$ de variables de Bernoulli indépendants ; alors

$X = X_1 + X_2 + \cdots + X_n$ . Donc $m_K(X) = E(X^k) = E(X_1^k + X_2^k + \cdots + X_n^k) = \sum_{i=1}^{n} E(X_i^k) = n.p$

$$E(X) = E(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} E(X_i) = n\,E(X_i) = n.p$$

De même la variance est égale à : $V(X) = V(X_1 + X_2 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) = n\,V(X_i) = n.p.q$

$$M_X(t) = E(e^{xt}) = E\left(e^{t(X_1+X_2+\cdots+X_n)}\right) = E(e^{tX_1})E(e^{tX_2}) \dots \dots \dots E(e^{tX_n}) = E(e^{tX_i})^n = (q + e^t p)^n$$

**3) Remarque :**

- Si $X \rightsquigarrow \beta(n; p)$ alors$(n - X) \rightsquigarrow \beta(n; 1 - p) = \beta(n; q)$.
- Si $X \rightsquigarrow \beta(n; p)$ et $Y \rightsquigarrow \beta(m; p)$ avec X et Y indépendants alors $X + Y \rightsquigarrow \beta(n + m; p)$ .

## Section 4 : **La loi de poisson:**

1) Définition :

La V.A qui suit la loi de poisson peut prendre un nombre variable de résultats de façon que la probabilité d'obtention d'une valeur k par la V.A X décroit très rapidement lorsque k augmente.

C'est la loi des évènements rares comme par exemple :

Le nombre d'appels téléphoniques reçus pendant une minute par un standard (la probabilité d'avoir un appel est supérieur à la probabilité d'avoir cinq appels durant une minute) ; le nombre de bateaux qui rentre au port pendant une journée : plus le nombre augmente plus la probabilité diminue.

X suit la loi de poisson de paramètre, notée $\lambda$ par : $X \rightsquigarrow \wp(\lambda)$ tel que : $P(X = x) = \begin{cases} e^{-\lambda} \frac{\lambda^x}{x!} & \text{si } k \in \mathbb{N} \\ 0 & \text{sinon} \end{cases}$

2) Caractéristiques :

$$E(X) = \sum_{x=0}^{\infty} k\,P(X = x) = \sum_{x=0}^{\infty} x\,e^{-\lambda} \frac{\lambda^x}{x!} = \lambda\,e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

Or $e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \Longrightarrow$ on pose k=x-1 alors si $x \longrightarrow 1$ on a $k \longrightarrow 0$ donc $\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \sum_{k=0}^{\infty} \frac{\lambda^x}{k!} = e^{\lambda}$

$$E(X) = \lambda\,e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda\,e^{-\lambda}\ e^{\lambda}\ = \lambda$$

De même nous avons : $V(X) = E(X) = \lambda$.

$$M_X(t) = E(e^{Xt}) = \sum_{x=0}^{\infty} e^{xt}\,P(X = x) = \sum_{x=0}^{\infty} e^{xt} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda}.e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

**Remarque :** Si $X \rightsquigarrow \wp(\lambda)$ et $Y \rightsquigarrow \wp(\theta)$ avec X et Y indépendants alors $X + Y \rightsquigarrow \wp(\lambda + \theta)$.

## Section 5 : **La loi Géométrique :**

1) Définition :

On répète des expériences de Bernoulli (P) jusqu'à l'observation du premier succès ; le nombre d'essais nécessaires à l'obtention de ce premier succès est une V.A $X \leadsto G(P)$.

On a pour tout entier positif $x$ : $P(X = x) = p(1-p)^{x-1} = pq^{x-1}$

2) Caractéristiques :

$$E(X) = \frac{1}{p} \text{ et } V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$$

# III- Les lois de probabilités continues :

## Section 1 : **La loi Uniforme :**

La loi de probabilité uniforme définie sur $[a\,,b]$ généralise la notion d'équiprobabilité au cas continue. Sa densité de probabilité est constante sur $[a\,,b]$ et vaut : $f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a\,,b] \\ 0 \text{ sinon} \end{cases}$

La fonction de répartition d'une V.A.C Uniforme est données par :

$$F(x) = \begin{cases} 0 \text{ si } x < a \\ \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{a} 0dt + \int_{a}^{x} \frac{1}{b-a}dt = \frac{x-a}{b-a} \text{ si } x \in [a\,,b] \\ 1 \text{ si } x > b \end{cases}$$

1) Caractéristiques :

$$m_k = E(X^k) = \int_a^b x^k \frac{1}{b-a}dt = \frac{1}{b-a}\left[\frac{x^{k+1}}{k+1}\right]_a^b = \frac{b^{k+1} - a^{k+1}}{(b-a)(k+1)}$$

$$m_1 = E(X) = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2} \text{ et } m_2 = E(X^2) = \frac{b^3 - a^3}{3(b-a)}$$

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

$$M_X(t) = E(e^{tx}) = \int_a^b e^{tx} \frac{1}{b-a} dx = \frac{e^{tb} - e^{ta}}{b-a}$$

## Section 2 : **La loi Gamma : $\gamma$**

1) Présentation de la fonction Gamma : $\gamma$

Soit la fonction Gamma définie comme suit : $\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt \quad telque \quad \alpha \in IN \text{ et } t \in IR^+$

Cette relation satisfait plusieurs relations utiles ; en particulier :

$\Gamma(\alpha+1) = \int_0^{+\infty} t^{\alpha} e^{-t} dt = \int_0^{+\infty} \alpha t^{\alpha-1} e^{-t} dt + [-t^{\alpha} e^{-t}]_0^{+\infty} = \alpha\Gamma(\alpha)$ si $\alpha > 0$

$\Gamma(1) = \int_0^{+\infty} t^{1-1} e^{-t} dt = [-e^{-t}]_0^{+\infty} = 1 \Longrightarrow \Gamma(1) = 1$ et $\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} t^{-\frac{1}{2}} e^{-t} dt = \sqrt{\pi}$

En se basant sur les deux formules précédentes ; on a :

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha) = \alpha(\alpha-1)\Gamma(\alpha-1) = \alpha(\alpha-1)(\alpha-2) \dots (2)\Gamma(1) = \alpha! \implies \Gamma(\alpha+1) = \alpha!$

**2)** La densité de probabilité (d.d.p) de la loi Gamma :

On dit qu'une V.A X suit la loi Gamma de paramètre $\alpha$ et $\beta$ ; notée : $X \rightsquigarrow \gamma(\alpha, \beta)$ si et seulement si :

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad si\ x > 0\ ;\ \alpha > 0\ \ et\ \ \beta > 0.$$

**3)** Caractéristiques :

a) **La fonction génératrice des moments et l'espérance mathématique**:

$$M_X(t) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{+\infty} e^{tx} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{+\infty} x^{\alpha-1} e^{-\left(\frac{1}{\beta}-t\right)x} dx = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\left(\frac{1}{\beta}-t\right)}} dx. \quad /1/$$

La d.d.p de la loi $\gamma$ est comme suit : $f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}$ $\qquad si\ x > 0\ ;\ \alpha > 0\ \ et\ \ \beta > 0.$

- On sait que : $\int_0^{+\infty} f(x)dx = 1 = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \Leftrightarrow \Gamma(\alpha)\beta^{\alpha} = \int_0^{+\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx. \quad /2/$

En remplaçant /2/ dans /1/ ; on trouve :

$M_X(t) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \Gamma(\alpha) \left(\frac{\beta}{(1-\beta t)}\right)^{\alpha} = \left(\frac{1}{1-\beta t}\right)^{\alpha} = (1-\beta t)^{-\alpha}\ \ si\ \ 1-\beta t > 0 \Leftrightarrow t < \frac{1}{\beta}.$

En effet si $t \geq \frac{1}{\beta}$ alors la quantité $\frac{1}{\beta} - t$ dans l'intégrale /1/ n'est pas positif et l'intégrale /2/ est alors infinie ; donc la loi Gamma n'existe que si $< \frac{1}{\beta}$.

$$E(X) = \frac{\partial M_X(t)}{\partial t}|_{t=0} = \frac{\partial(1-\beta t)^{-\alpha}}{\partial t}|_{t=0} = \frac{\alpha.\beta}{(1-\beta t)^{\alpha+1}}|_{t=0} = \alpha.\beta$$

b) **Le moment non centré d'ordre K et la variance:**

$$m_K(X) = E(X^K) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \int_0^{+\infty} x^{(K+\alpha)-1} e^{-\frac{x}{\beta}} dx = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \Gamma(\alpha+K)\beta^{\alpha+K} = \frac{\Gamma(\alpha+K)\beta^K}{\Gamma(\alpha)}$$

Or La fonction $Z = x^{(K+\alpha)-1} e^{-\frac{x}{\beta}}$ est la densité de la loi Gamma tel que $Z \rightsquigarrow \gamma(K+\alpha; \beta)$ donc

$$\frac{1}{\Gamma(K+\alpha)\beta^{K+\alpha}} \int_0^{+\infty} x^{(K+\alpha)-1} e^{-\frac{x}{\beta}} dx = 1 \Leftrightarrow \int_0^{+\infty} x^{(K+\alpha)-1} e^{-\frac{x}{\beta}} dx = \Gamma(K+\alpha)\beta^{K+\alpha}$$

De même on a : $V(X) = E(X^2) - E(X)^2 = \frac{\Gamma(\alpha+2)\beta^2}{\Gamma(\alpha)} - (\alpha.\beta)^2 = \alpha(\alpha+1).\beta^2 - (\alpha.\beta)^2 = \alpha.\beta^2$

**4)** Théorème de stabilité de la loi Gamma :

Soit X et Y deux V.A indépendants suivant respectivement $\gamma(p;\theta)$ $et$ $\gamma(q\ ;\theta)$ alors la V.A

$Z = X + Y \ \backsim\ \gamma(p + q;\theta)$.

## Section 3 : **La loi Exponentielle :**

La loi exponentielle est très utilisée pour étudier la durée de vie et en informatique pour modéliser la durée du temps. Sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \theta e^{-\theta x} & \text{si } x \geq 0 \text{ et } \theta > 0 \\ 0 & \text{ailleurs} \end{cases}$$

On l'écrit $X \backsim \exp(\theta)$

On remarque la densité de la loi exponentielle de paramètre $\theta > 0$ est identique à la loi Gamma de paramètre$\left(1;\frac{1}{\theta}\right)$ donc $X \backsim \exp(\theta) \Leftrightarrow X \backsim \gamma(1,\frac{1}{\theta})$

**1)** Caractéristiques :

La fonction de répartition de cette loi est donnée par : Soit $x \geq 0$ et $\theta > 0$.

$$F(x) = P(X \leq x) = \int_0^x \theta e^{-\theta t}dt = 1 - e^{-\theta x}$$

Les caractéristiques la loi exponentielle $X \backsim \exp(\theta)$sont déduites directement de celles de la loi Gamma : $X \backsim \gamma(1,\frac{1}{\theta})$.

$$M_X(t) = E(e^{Xt}) = \left(1 - \frac{t}{\theta}\right)^{-1} = \frac{\theta}{\theta - t} \quad t < \theta$$

$$m_K(X) = E(X^K) = \frac{\Gamma(K+1)\left(\frac{1}{\theta}\right)^K}{\Gamma(1)} = \Gamma(K+1)\left(\frac{1}{\theta}\right)^K = (K!)\left(\frac{1}{\theta}\right)^K$$

$$E(X) = 1 \times \frac{1}{\theta} = \frac{1}{\theta} \qquad et \qquad V(X) = 1 \times \left(\frac{1}{\theta}\right)^2 = \frac{1}{\theta^2}$$

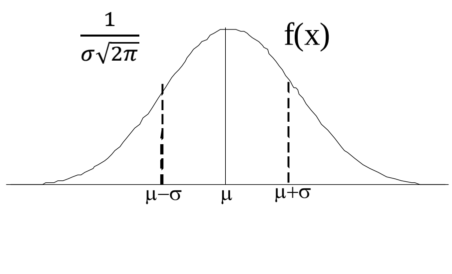
## Section 4 : **La loi Normal (ou de Laplace – Gauss):**

Cette loi joue un rôle central dans les études statistiques. En effet, la loi Normale ainsi que les lois qui lui sont associés sont assez maniable analytiquement.

On dit qu'une V.A X suit la loi Normale de paramètres $\mu,\sigma$ et on la note : $X \backsim N(\mu,\sigma)$. Sa d.d.p est :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\ e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad pour\ tout\ x \in IR$$

**1)** Représentation et propriétés de la d.d.p :



La d.d.p est une fonction symétrique par rapport à $x = \mu$ ; elle atteint son maximum au point $\left(\mu, \frac{1}{\sigma\sqrt{2\pi}}\right)$ et elle admet deux points d'inflexion : $x_1 = \mu - \sigma \quad et \quad x_2 = \mu + \sigma$.

La fonction de répartition de la loi Normale est de la forme :

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

En se basant sur la propriété de la symétrie : $F(x) = 1 - F(-x)$

$$P(|X| \leq x) = P(-x \leq X \leq x) = F(x) - F(-x) = F(x) - [1 - F(x)] = 2F(x) - 1$$

**2)** <u>Caractéristiques :</u>

$$M_X(t) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{xt} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \; on \, pose \; z = \frac{x-\mu}{\sigma} \Leftrightarrow x = z\sigma + \mu \Longrightarrow dz = \frac{dx}{\sigma}$$

Bornes : si $x \to +\infty \quad alors \, z \to +\infty \quad et \quad si \, x \to -\infty \quad alors \, z \to -\infty$

$$M_X(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} \sigma \, e^{-\frac{z^2}{2}} \; e^{t(z\sigma+\mu)} dz = \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(z^2 - 2tz\sigma)} dz$$

Or $z^2 - 2tz\sigma = (z - t\sigma)^2 - t^2\sigma^2 \Longrightarrow M_X(t) = \frac{e^{\frac{1}{2}t^2\sigma^2} e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(z-t\sigma)^2} dz$

Soit $y = z - t\sigma \quad \Longrightarrow \quad dy = dz$

$M_X(t) = \frac{e^{\frac{1}{2}t^2\sigma^2 + t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy$ or dans le cas où $\mu = 0 \; et \; \sigma = 1$ ; la loi Normale à pour d.d.p la fonction $f(y) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{y^2}{2}}$ et on sait que $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \; e^{-\frac{y^2}{2}} \, dy = 1$ alors $M_X(t) = e^{\frac{1}{2}t^2\sigma^2 + t\mu}$

$$E(X) = M_X'(t)|t_{=0} = (t\sigma^2 + \mu)e^{\frac{1}{2}t^2\sigma^2 + t\mu}|t_{=0} = \mu$$

$$E(X^2) = M_X''(t)|t_{=0} = \left[\sigma^2 e^{\frac{1}{2}t^2\sigma^2 + t\mu} + (\mathrm{t}\sigma^2 + \mu)e^{\frac{1}{2}t^2\sigma^2 + t\mu}\right]|t_{=0} = \sigma^2 + \mu^2$$

$$V(X) = (\sigma^2 + \mu^2) - \mu^2 = \sigma^2$$

Donc la loi Normale est définie par sa moyenne $\mu$ et son écart type $\sigma$.

## <u>Section 6 :</u> **La loi Normal centrée réduite :**

On dit qu'une V.A X est Normale centrée réduite si elle a comme moyenne 0 et comme écart type 1 ; notée N(0 ;1). Sa d.d.p s'écrit comme suit : $\phi(x) = \frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} \quad pour \, tout \, x \in IR$

On remarque que $\phi(x)$ est continue donc on a : $\phi'(x) = \frac{-x}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}} = -x\phi(x)$ et $\phi''(x) = \frac{(x^2-1)}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}}$

Sa fonction de répartition s'écrit alors comme suit :

$$\Phi(x) = P(X \leq x) = \int_{-\infty}^{x} \phi(t)dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \; e^{-\frac{t^2}{2}}dt$$

On sait que : $\Phi(x) = 1 - \Phi(-x) \quad et \quad P(|X| \leq x) = 2\Phi(x) - 1$

Puisque la loi Normale centrée réduite est symétrique par rapport à l'origine alors elle atteint son maximum au point $(0;\frac{1}{\sqrt{2\pi}})$ et elle possède deux points d'inflexions : $x_1 = 1 \;\; et \;\; x_2 = -1$.

**1)** Caractéristiques :

$$E(X) = \int_{-\infty}^{+\infty} x\phi(x)dx = \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\left[\; e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} = 0$$

$$V(X) = E(x^2) = \int_{-\infty}^{+\infty} x^2\phi(x)dx = \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}}dx = \frac{1}{\sqrt{2\pi}}\left[-x \; e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \; e^{-\frac{x^2}{2}}dx = 1$$

Avec $g' = x \; e^{-\frac{x^2}{2}} \qquad et\; h = x$

$$M_X(t) = \; e^{-\frac{t^2}{2}}$$

**2)** Théorème de stabilité de la loi Normale :

Toute combinaison affine de V.A Normales indépendantes est une V.A Normale. Par exemple : $\boldsymbol{X_1 \leadsto N(\mu_1, \sigma_1)}$ ; $X_2 \leadsto N(\mu_2, \sigma_2)$ et $X_1;\; X_2$ sont deux V.A indépendantes alors pour tout V.A Y on a : $Y = aX_1 + b\,X_2$ avec $a$ et $b$ sont des constantes quelconques : $Y \leadsto N(a\mu_1 + b\mu_2; \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$

**a) Démonstration :**

$$M_Y(t) = M_{X_1}(t)M_{X_2}(t) = \left(e^{\frac{1}{2}t^2\sigma_1^2 + t\mu_1}\right)\left(e^{\frac{1}{2}t^2\sigma_2^2 + t\mu_2}\right) = e^{\frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)+t(\mu_1+\mu_2)}$$

Donc $Y \leadsto N(a\mu_1 + b\mu_2; \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2})$

**3)** Lecture de la table de la loi Normale :

Pour pouvoir utiliser la table, il faut d'abord centrée et réduire $X \leadsto N(\mu, \sigma) \Longrightarrow Z = \frac{X-\mu}{\sigma} \leadsto N(0{,}1)$

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

- Dans le cas où $x < 0$ ; on la traite en utilisant la relation $\Phi(x) = 1 - \Phi(-x)$
- La probabilité se trouve à l'intérieur de la table, par contre la valeur de x est la somme de deux valeurs ; l'une se trouvant sur la première ligne de la table et l'autre sur sa première colonne.

a) **Applications :**

1/ Déterminer $\Phi(1{,}96)$ ; $P(-1{,}65 < X < 0{,}51)$ et $P(|X| < 1{,}18)$ et aussi $\Phi(x) = 0{,}9842$
2/ Déterminer $P(X < 7)$ et $P(2{,}5 < X < 3)$ tel que $X \leadsto N(3{,}2)$

b) **Réponse :**

# IV- Les lois dérivées de la loi Normale:

Section 1 : **La loi de Khi- Deux : $\chi_n^2$**

Soient $X_1, X_2, \dots\dots\dots X_n$ une de n V.A i.i.d distribuées suivant la loi Normale centrée réduite.

On pose $Y = \chi_n^2 = X_1^2 + X_2^2 + \cdots \dots \dots + X_n^2$ suit une loi de probabilité appelée: Khi- Deux$(\chi_n^2)$ de degrés de libertés n ; sa d.d.p est de la forme :

$$f(y) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} & si \;\; y > 0 \\ 0 & sinon \end{cases}$$

$Si\ n = 1\ on\ a :\ f(y) = \dfrac{1}{\sqrt{2\pi}} \dfrac{e^{-\frac{y}{2}}}{\sqrt{y}} = \phi(\sqrt{y}) y^{-\frac{1}{2}}$

**1)** Caractéristiques :

On remarque la densité de la loi de Khi- Deux de paramètre est identique à la loi Gamma de paramètre$\left(\frac{n}{2}; 2\right)$ donc $X \leadsto \chi_n^2 \Leftrightarrow X \leadsto \gamma(\frac{n}{2}; 2)$

Les caractéristiques la loi de Khi- Deux sont déduites directement de celles de la loi Gamma tel que $X \leadsto \gamma(\frac{n}{2}, 2)$.

$$M_X(t) = (1 - 2t)^{-\frac{n}{2}} \;\; si \;\; 1 - 2t > 0 \;\Leftrightarrow t < \frac{1}{2}.$$

$$E(X) = \frac{n}{2} \times 2 = n$$

$$V(X) = \frac{n}{2} \times 2^2 = 2n$$

$$m_K(X) = E(X^K) = \frac{(2^K) \times \Gamma\left(\frac{n}{2} + K\right)}{\Gamma\left(\frac{n}{2}\right)}$$

**2)** Stabilité de la loi de Khi- Deux :

La loi de Khi- Deux est stable par addition ; en effet, si $Y_1 \leadsto \chi_{n_1}^2\ et\ Y_2 \leadsto \chi_{n_2}^2$ alors $Y_1 + Y_2 \leadsto \chi_{n_1+n_2}^2$

**3)** Lecture de la table :

La première ligne représente les probabilités ; la première colonne représente les degrés de libertés et à l'intérieur de cette table on retrouve les valeurs prisent par la variable aléatoire Y.

a) **Exemples :**

Soient $Y_1 \leadsto \chi^2_{10}$ ; $Y_2 \leadsto \chi^2_{19}$ et $Y = Y_1 + Y_2$ ; chercher $a$ si $P(Y_1 > a) = 0{,}99$ et calculer $P(Y_2 < 21{,}689)$
Chercher b tel que : $P(Y > b) = 0{,}25$

**b) Réponses :**

## Section 2 : La loi de Student :

Soit $(X,Y)$ un couple de V.A indépendantes tel que $X \leadsto N(0{,}1)$ et $Y \leadsto \chi^2_n$ alors la V.A $T = \frac{X}{\sqrt{\frac{Y}{n}}}$ possède une densité de probabilité : $f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)} \times \left(n^{-\frac{1}{2}}\right) \times \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$ ; $\forall\, t \in IR$

La loi de probabilité de t est appelée loi de Student de n degré de liberté ; noté par : $T \leadsto t_{(n)}$.

**Remarque :** La distribution de Student est une loi symétrique càd $P(t \le a) = 1 - P(t \le -a) \Leftrightarrow F(a) = 1 - F(-a)$.

**1)** Caractéristiques :

Comme $X \leadsto N(0{,}1)$ et $T = \frac{X}{\sqrt{\frac{Y}{n}}} = X.\frac{1}{\sqrt{\frac{Y}{n}}}$ , on a alors :

$$E(T) = E\left(X.\frac{1}{\sqrt{\frac{Y}{n}}}\right) = E(X).E\left(\frac{1}{\sqrt{\frac{Y}{n}}}\right) = 0 \; car \; E(X) = 0 \; donc \; E(T) = 0 \; si \; n > 1$$

$$V(T) = E(T^2) - (E(T))^2 = E(X^2).E\left(\left(\frac{1}{\sqrt{\frac{Y}{n}}}\right)^2\right) = E\left(\frac{1}{\frac{Y}{n}}\right) = \frac{n}{n-2} \; si \; n > 2.$$

**2)** Lecture de la table :

La première ligne représente les probabilités ; la première colonne représente les degrés de libertés et à l'intérieur de cette table on retrouve les valeurs prisent par la variable aléatoire T.

a) **Exemples :** soient $T \leadsto t_{(20)}$ ; chercher a tel que $P(T \ge a) = 0{,}2$ et $P(T < a) = 0{,}01$

**b) Réponses :**

## Section 3 : La loi de Fisher- Snedecor :

Soit $(X,Y)$ un couple de V.A indépendantes tel que $X \leadsto \chi^2_{n_1}$ et $Y \leadsto \chi^2_{n_2}$ alors la V.A $W = \frac{X/n_1}{Y/n_2}$ suit la loi de Fisher de degrés de libertés $n_1$ et $n_2$.

Cette loi admet comme d.d.p :

$$f(w) = \begin{cases} 0 \quad si \quad w \leq 0 \\ \dfrac{\Gamma(\frac{n_1+n_2}{2}) n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{\Gamma\left(\frac{n_1}{2}\right)\Gamma(\frac{n_2}{2})} \times \dfrac{w^{\frac{n_1}{2}-1}}{(n_2 + n_1 \mathrm{w})^{\left(\frac{n_1+n_2}{2}\right)}} \end{cases}$$

**1)** Caractéristiques :

$$E(W) = \frac{n_2}{n_2 - 2} \ si \ \ n_2 > 2 \ \ et \, V(W) = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(\mathrm{n}_2 - 2)^2(n_2 - 4)} \quad si \, n_2 > 4$$

**a) Remarque :** Cas où ($n_1$) ou ($n_2$) ne figure pas sur la table :

$$si \, W \rightsquigarrow F(n_1; n_2) \Leftrightarrow W = \frac{\chi^2_{n_1}/_{n_1}}{\chi^2_{n_2}/_{n_2}} \Leftrightarrow \frac{1}{W} = \frac{\chi^2_{n_2}/_{n_2}}{\chi^2_{n_1}/_{n_1}} \Leftrightarrow \frac{1}{W} \rightsquigarrow F(n_2; n_1)$$

**2)** Lecture de la table :

C'est une table à double entrée tel que dans la première ligne nous avons le premier degré de liberté ($n_1$) et dans la première colonne nous avons le second degré de liberté ($n_2$) et à l'intérieur de cette table on retrouve les valeurs prisent par la variable aléatoire W pour les probabilités 0,95 ;0,975 ;0,99……….

**3)** Applications :

1/ soit $W \rightsquigarrow F(7; 26)$ chercher a tel que : $P(W \leq a) = 0{,}95$ et $P(W > a) = 0{,}025$

2/ soit $W \rightsquigarrow F(24; 1)$ .Chercher$P(W \leq a) = 0{,}025$.

**4)** Réponse :

# V- Approximation :

**1)** Définition :

1) Pour n grand la loi Binomiale peut être approximé soit par la loi de Poisson (si P est petite) ; soit par la loi Normale ; en d'autre terme si :

$- \, si \, n \geq 30 \, et \, p \leq 0{,}1 \ \ c\grave{a}d \, npq \leq 10 \, alors \, B(n, p)$ est approximé à la loi de Poisson $\wp(\lambda = np)$.

$- \, si \, n > 30$ ; $np > 15 \, et \, npq > 15 \ \ alors \, B(n, p)$ est approximé à la loi Normale $N(\mu, \sigma)$ tel que $\mu = np$ et $\sigma = \sqrt{np(1-p)} = \sqrt{npq}$.

2) Lorsque $\lambda$ est grand càd $\lambda \geq 10$ la loi de Poisson est approximé à la loi Normale de paramètre $\mu = \lambda$ et $\sigma = \sqrt{\lambda}$.

## 2) Applications :

### a) Exercice 1 :

Dans une chaine de fabrication, 5% des pièces sont défectueuses ; on prélève une pièce, on examine si elle est défectueuse et on la replace parmi les autres. On répète 120 fois cette expérience. On désigne par X la variable aléatoire qui à chaque tirage de 120 pièces associe le nombre des pièces défectueuses.
1/ Identifier la loi suivie par X.
2/ Donner une approximation à cette loi.

### b) Réponse 1:

### c) Exercice 2 :

On jette un dé 180 fois. On note par $X$ la variable aléatoire : « le nombre obtenue est 3 ».
1/ Quelle est la loi de $X$ ?
2/ Donner une approximation à cette loi.
3/ Calculez la probabilité pour que X soit compris entre 29 et 32.

### d) Réponse 2 :