

Centre universitaire Abdelhafid Boussouf -Mila  
Institut de Sciences et Technologies  
Département de mathématiques et informatique  
3<sup>ème</sup> Année mathématiques

# **Matière: Introduction aux probabilités et statistique descriptive**

## **Chapitre 1: Notions de base et vocabulaire statistique**

Présentée par:  
AZI Mourad

March 6, 2023



## Introduction



## Introduction

### Concepts de base de la statistique

Population et unité statistique

Variabes statistiques



## Introduction

### Concepts de base de la statistique

- Population et unité statistique

- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète

- Variable statistique continue



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### Variable quantitative discrète: diagramme en bâtons

#### Variabes continues ou réparties en classes: Histogramme

#### Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue



Ce premier chapitre donne les définitions et les propriétés des principales notions utiles pour comprendre et traiter un problème de statistique.



Ce premier chapitre donne les définitions et les propriétés des principales notions utiles pour comprendre et traiter un problème de statistique.

La statistique est la discipline qui **étudie des phénomènes à travers la collecte de données**. Elle s'intéresse à l'analyse, à l'interprétation, à la présentation et à l'utilisation des données, afin de les rendre compréhensibles par tous et prendre des décisions et résoudre des problèmes.



Ce premier chapitre donne les définitions et les propriétés des principales notions utiles pour comprendre et traiter un problème de statistique.

La statistique est la discipline qui étudie des phénomènes à travers la collecte de données. Elle s'intéresse à **l'analyse, à l'interprétation, à la présentation et à l'utilisation des données**, afin de les rendre compréhensibles par tous et prendre des décisions et résoudre des problèmes.





Ce premier chapitre donne les définitions et les propriétés des principales notions utiles pour comprendre et traiter un problème de statistique.

La statistique est la discipline qui étudie des phénomènes à travers la collecte de données. Elle s'intéresse à l'analyse, à l'interprétation, à la présentation et à l'utilisation des données, afin de les rendre **compréhensibles** par tous et **prendre des décisions et résoudre des problèmes**.



## Historique

L'histoire des statistiques remonte à l'antiquité, où les données étaient collectées pour les besoins de recensement et d'impôts.

- ▶ 17ème siècle, des statisticiens tels que John Graunt en Angleterre ont utilisé des méthodes statistiques pour étudier les populations et les démographies;



## Historique

L'histoire des statistiques remonte à l'antiquité, où les données étaient collectées pour les besoins de recensement et d'impôts.

- ▶ 17ème siècle, des statisticiens tels que John Graunt en Angleterre ont utilisé des méthodes statistiques pour étudier les populations et les démographies;
- ▶ 18ème siècle, les mathématiciens tels que Pierre-Simon Laplace ont utilisé des méthodes statistiques pour étudier les phénomènes naturels et les lois de la nature;



## Historique

L'histoire des statistiques remonte à l'antiquité, où les données étaient collectées pour les besoins de recensement et d'impôts.

- ▶ 17ème siècle, des statisticiens tels que John Graunt en Angleterre ont utilisé des méthodes statistiques pour étudier les populations et les démographies;
- ▶ 18ème siècle, les mathématiciens tels que Pierre-Simon Laplace ont utilisé des méthodes statistiques pour étudier les phénomènes naturels et les lois de la nature;
- ▶ 19ème siècle, les statisticiens tels que Carl Friedrich Gauss et Adolphe Quetelet ont développé des méthodes statistiques pour étudier les relations entre les variables et les phénomènes sociaux;



## Historique

L'histoire des statistiques remonte à l'antiquité, où les données étaient collectées pour les besoins de recensement et d'impôts.

- ▶ 17ème siècle, des statisticiens tels que John Graunt en Angleterre ont utilisé des méthodes statistiques pour étudier les populations et les démographies;
- ▶ 18ème siècle, les mathématiciens tels que Pierre-Simon Laplace ont utilisé des méthodes statistiques pour étudier les phénomènes naturels et les lois de la nature;
- ▶ 19ème siècle, les statisticiens tels que Carl Friedrich Gauss et Adolphe Quetelet ont développé des méthodes statistiques pour étudier les relations entre les variables et les phénomènes sociaux;
- ▶ Au début du 20ème siècle, les statisticiens tels que Ronald A. Fisher et Jerzy Neyman ont élaboré de nouvelles méthodes statistiques pour les expériences scientifiques et les tests de signification.



## Domaine d'application

Les statistiques sont largement utilisées dans de nombreux domaines, Elles permettent une meilleure compréhension des phénomènes et une prise de décision plus informée:

- ▶ **Sciences sociales : analyse des données démographiques, étude des comportements sociaux et économiques;**



## Domaine d'application

Les statistiques sont largement utilisées dans de nombreux domaines, Elles permettent une meilleure compréhension des phénomènes et une prise de décision plus informée:

- ▶ Sciences sociales : analyse des données démographiques, étude des comportements sociaux et économiques;
- ▶ Sciences économiques : **prévision des tendances économiques, analyse des données financières et économiques, étude de la satisfaction des clients et analyse de la consommation de produits;**



## Domaine d'application

Les statistiques sont largement utilisées dans de nombreux domaines, Elles permettent une meilleure compréhension des phénomènes et une prise de décision plus informée:

- ▶ Sciences sociales : analyse des données démographiques, étude des comportements sociaux et économiques;
- ▶ Sciences économiques : prévision des tendances économiques, analyse des données financières et économiques, étude de la satisfaction des clients et analyse de la consommation de produits;
- ▶ **Sciences informatiques : traitement de grandes quantités de données, reconnaissance de la voix et des images et traitement automatique des langues naturelles;**





## Domaine d'application

Les statistiques sont largement utilisées dans de nombreux domaines, Elles permettent une meilleure compréhension des phénomènes et une prise de décision plus informée:

- ▶ Sciences sociales : analyse des données démographiques, étude des comportements sociaux et économiques;
- ▶ Sciences économiques : prévision des tendances économiques, analyse des données financières et économiques, étude de la satisfaction des clients et analyse de la consommation de produits;
- ▶ Sciences informatiques : traitement de grandes quantités de données, reconnaissance de la voix et des images et traitement automatique des langues naturelles;
- ▶ **Sciences de l'environnement : étude des changements climatiques, analyse de la qualité de l'air et de l'eau;**



## Domaine d'application

- ▶ **Physique:** Les statistiques peuvent être utilisées pour traiter les données issues d'expériences et d'observations pour déterminer des lois et des corrélations quantitatives, évaluer les incertitudes et les erreurs des mesures, pour effectuer des tests de signification et pour déterminer les intervalles de confiance;



## Domaine d'application

- ▶ Physique: Les statistiques peuvent être utilisées pour traiter les données issues d'expériences et d'observations pour déterminer des lois et des corrélations quantitatives, évaluer les incertitudes et les erreurs des mesures, pour effectuer des tests de signification et pour déterminer les intervalles de confiance;
- ▶ Sports : analyse des performances athlétiques, évaluation des stratégies d'entraînement;



## Domaine d'application

- ▶ Physique: Les statistiques peuvent être utilisées pour traiter les données issues d'expériences et d'observations pour déterminer des lois et des corrélations quantitatives, évaluer les incertitudes et les erreurs des mesures, pour effectuer des tests de signification et pour déterminer les intervalles de confiance;
- ▶ Sports : analyse des performances athlétiques, évaluation des stratégies d'entraînement;
- ▶ **Éducation : évaluation des programmes d'enseignement, analyse des performances des élèves.**



## But des statistiques descriptives

La statistique descriptive a pour but :

- ▶ **structurer et de représenter l'information contenue dans les données;**



## But des statistiques descriptives

La statistique descriptive a pour but :

- ▶ structurer et de représenter l'information contenue dans les données;
- ▶ de dégager les propriétés essentielles que l'on peut déduire d'une accumulation de données ;



## But des statistiques descriptives

La statistique descriptive a pour but :

- ▶ structurer et de représenter l'information contenue dans les données;
- ▶ de dégager les propriétés essentielles que l'on peut déduire d'une accumulation de données ;
- ▶ **de donner une image concise et simplifiée de la réalité.**



## Introduction

### Concepts de base de la statistique

Population et unité statistique

Variabes statistiques

### Rangement des données (tableau statistique)

Variable statistique discrète

Variable statistique continue

### Les différents modes de représentation graphique des données

Variable Qualitative

Diagramme en tuyaux d'orgues(Barre)

Diagramme circulaire

Variable quantitative discrète: diagramme en bâtons

Variabes continues ou réparties en classes: Histogramme

Graphiques de fréquences cumulées, diagramme intégral

Variable discrète

Cas d'une variable statistique continue





## Définition 1 (Ensemble statistique ou population)

Désigne un ensemble d'individus (unités statistiques) sur lesquels on étudie une ou plusieurs propriétés. Cet ensemble est noté  $\omega$ .



## Définition 1 (Ensemble statistique ou population)

Désigne un ensemble d'individus (unités statistiques) sur lesquels on étudie une ou plusieurs propriétés. Cet ensemble est noté  $\omega$ .

## Définition 2 (Unité statistique)

Les éléments qui composent une population statistique sont appelés unités statistiques, dans le langage des statisticiens, On parle aussi d'individus (les premières statistiques étant des études démographiques). l'unité statistique peut être un individu, un ménage, une entreprise, un établissement, une commune, un département, une région ou encore un pays.

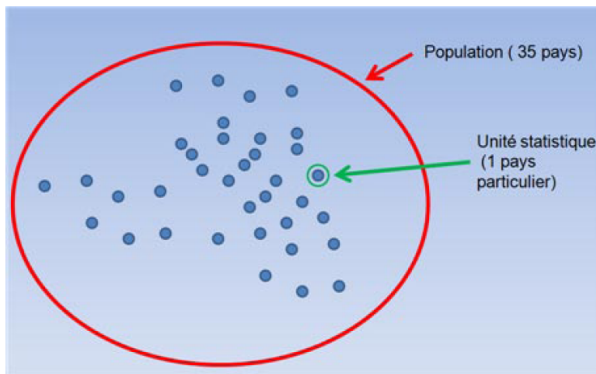


Figure: Population, Individu



## Introduction

### Concepts de base de la statistique

Population et unité statistique

**Variabes statistiques**

### Rangement des données (tableau statistique)

Variable statistique discrète

Variable statistique continue

### Les différents modes de représentation graphique des données

Variable Qualitative

Diagramme en tuyaux d'orgues(Barre)

Diagramme circulaire

Variable quantitative discrète: diagramme en bâtons

Variabes continues ou réparties en classes: Histogramme

Graphiques de fréquences cumulées, diagramme intégral

Variable discrète

Cas d'une variable statistique continue



## Caractères

C'est ce qui est observé ou mesuré sur les individus d'une population statistique. Il peut s'agir d'un caractère

- ▶ **quantitatif;**



## Caractères

C'est ce qui est observé ou mesuré sur les individus d'une population statistique. Il peut s'agir d'un caractère

- ▶ quantitatif;
- ▶ **qualitatif.**



## Caractères

C'est ce qui est observé ou mesuré sur les individus d'une population statistique. Il peut s'agir d'un caractère

- ▶ quantitatif;
- ▶ qualitatif.

## Modalité

Une modalité est la valeur prise par une variable statistique qu'elle soit qualitative ou quantitative. Les modalités correspondent donc à l'ensemble des valeurs possibles.



## Caractères

C'est ce qui est observé ou mesuré sur les individus d'une population statistique. Il peut s'agir d'un caractère

- ▶ quantitatif;
- ▶ qualitatif.

## Modalité

Une modalité est la valeur prise par une variable statistique qu'elle soit qualitative ou quantitative. Les modalités correspondent donc à l'ensemble des valeurs possibles.

Si la Variable est " situation familiale " alors:

Les modalités sont " célibataire, marié, divorcé "





## Caractères quantitatifs

Sont les caractères qui sont représentés par des nombres et sur lesquels les opérations arithmétiques de base ont un sens, les modalités sont mesurables. Tels que : l'âge, la taille ou le poids d'un individu, nombre d'enfants...



## Caractères quantitatifs

Sont les caractères qui sont représentés par des nombres et sur lesquels les opérations arithmétiques de base ont un sens, les modalités sont mesurables. Tels que : l'âge, la taille ou le poids d'un individu, nombre d'enfants...

## Caractères qualitatifs

Sont tous les caractères qui ne sont pas représentés par des nombres, Ou qui sont parfois codés par des nombres sur lesquels les opérations arithmétiques de base n'ont pas de sens (ou un sens très limité). Les valeurs (les modalités) ne sont pas mesurable.



## Caractères quantitatifs

Une variable quantitative peut être une variable statistique discrète ou continue.

- ▶ Les variables statistiques discrètes: sont des variables qui ne peuvent prendre que des valeurs isolées, discrètes. Le nombre d'enfants d'une famille, le nombre de pétales d'une fleur, le nombre de buts marqués lors d'une rencontre de football... sont des variables quantitatives discrètes. Le plus fréquemment, les valeurs possibles sont des nombres entiers.



## Caractères quantitatifs

Une variable quantitative peut être une variable statistique discrète ou continue.

- ▶ Les variables statistiques discrètes: sont des variables qui ne peuvent prendre que des valeurs isolées, discrètes. Le nombre d'enfants d'une famille, le nombre de pétales d'une fleur, le nombre de buts marqués lors d'une rencontre de football... sont des variables quantitatives discrètes. Le plus fréquemment, les valeurs possibles sont des nombres entiers.
- ▶ Les variables statistiques continues peuvent prendre toutes les valeurs numériques possibles d'un ensemble inclus dans  $\mathbb{R}$ : le revenu, la taille, le taux de natalité sont des variables continues.

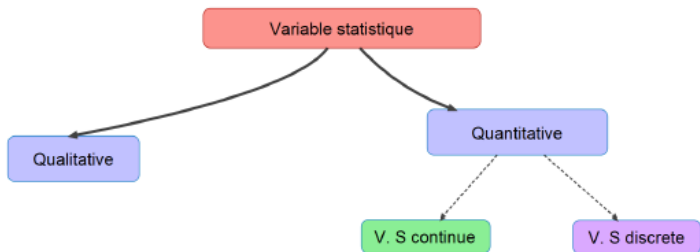


Figure: Type de variables statistique



## Tableau statistique

Un tableau permet une présentation synthétique des informations recueillies. D'une façon générale et selon que la variable est qualitative ou quantitative, discrète ou continue, un tableau se compose :

- ▶ D'une colonne indiquant les différentes modalités de la variable  $x_i$  ;
- ▶ D'une ou plusieurs autres colonnes indiquant l'effectif correspondant à ces diverses modalités, ainsi que les différentes fréquences.



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### Variable quantitative discrète: diagramme en bâtons

#### Variabes continues ou réparties en classes: Histogramme

#### Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue



## Distribution statistique discrète

Valeurs de la variable	Fréquences absolues	Fréquences relatives	Fréquences cumulées absolues	Fréquences cumulées relatives
$x_i$	$n_i$	$f_i$	$N_i$	$F_i = \sum_{k=1}^i f_k$

Soit une série statistique  $(x_i, n_i)_{i=1, \dots, p}$ . Nous appelons

- ▶ effectif  $n_i$  (Fréquences absolues) de la valeur  $x_i$  : le nombre de fois que la valeur  $x_i$  est prise;





## Distribution statistique discrète

Valeurs de la variable	Fréquences absolues	Fréquences relatives	Fréquences cumulées absolues	Fréquences cumulées relatives
$x_i$	$n_i$	$f_i$	$N_i$	$F_i = \sum_{k=1}^i f_k$

Soit une série statistique  $(x_i, n_i)_{i=1, \dots, p}$ . Nous appelons

- ▶ effectif  $n_i$  (Fréquences absolues) de la valeur  $x_i$  : le nombre de fois que la valeur  $x_i$  est prise;
- ▶ fréquence (Fréquences relatives) de la valeur  $x_i$ : le rapport de l'effectif de  $x_i$  à l'effectif total  $N$  de la population.  $f_i = \frac{n_i}{N}$ ;



## Distribution statistique discrète

Valeurs de la variable	Fréquences absolues	Fréquences relatives	Fréquences cumulées absolues	Fréquences cumulées relatives
$x_i$	$n_i$	$f_i$	$N_i$	$F_i = \sum_{k=1}^i f_k$

Soit une série statistique  $(x_i, n_i)_{i=1, \dots, p}$ . Nous appelons

- ▶ effectif  $n_i$  (Fréquences absolues) de la valeur  $x_i$  : le nombre de fois que la valeur  $x_i$  est prise;
- ▶ fréquence (Fréquences relatives) de la valeur  $x_i$ : le rapport de l'effectif de  $x_i$  à l'effectif total  $N$  de la population.  $f_i = \frac{n_i}{N}$ ;
- ▶ effectif cumulé en  $x_i$ :  $N_i = \sum_{j=1}^i n_j$ ;



## Distribution statistique discrète

Valeurs de la variable	Fréquences absolues	Fréquences relatives	Fréquences cumulées absolues	Fréquences cumulées relatives
$x_i$	$n_i$	$f_i$	$N_i$	$F_i = \sum_{k=1}^i f_k$

Soit une série statistique  $(x_i, n_i)_{i=1, \dots, p}$ . Nous appelons

- ▶ effectif  $n_i$  (Fréquences absolues) de la valeur  $x_i$  : le nombre de fois que la valeur  $x_i$  est prise;
- ▶ fréquence (Fréquences relatives) de la valeur  $x_i$ : le rapport de l'effectif de  $x_i$  à l'effectif total  $N$  de la population.  $f_i = \frac{n_i}{N}$ ;
- ▶ effectif cumulé en  $x_i$ :  $N_i = \sum_{j=1}^i n_j$ ;
- ▶ fréquence cumulée en  $x_i$  :  $F_i = \sum_{j=1}^i f_j$ ;



## Exemple 1 (variable statistique discrète)

Étude statistique du nombre de défauts sur une pièce de tissu.

Nombre de défauts	$n_i$	$f_i$	$N_i$	$F_i$
0	38	0,506	38	0,506
1	15	0,20	$53 = 38 + 15$	0,706
2	11	0,146	$64 = 53 + 11$	0,853
3	6	0,08	$70 = 64 + 6$	0,933
4	3	0,04	$73 = 70 + 3$	0,973
5	2	0,026	$75 = 73 + 2$	1



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue**

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### Variable quantitative discrète: diagramme en bâtons

#### Variabes continues ou réparties en classes: Histogramme

#### Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue



## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- Nombre de classe on peut le calculer avec la formul de Sturge:

$$K = 1 + 3.332 \log_{10}(N);$$



## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- ▶ Nombre de classe on peut le calculer avec la formul de Sturge:  
$$K = 1 + 3.332 \log_{10}(N);$$
- ▶ l'étendu de X:  $e = x_{max} - x_{min};$



## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- ▶ Nombre de classe on peut le calculer avec la formul de Sturge:  
$$K = 1 + 3.332 \log_{10}(N);$$
- ▶ l'étendu de X:  $e = x_{max} - x_{min};$
- ▶ l'amplitude d'une classe:  $a_i = e_i - e_{i-1}$ , on peut le calculer avec la formule:  
$$a_i = \frac{x_{max} - x_{min}}{k};$$





## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- ▶ Nombre de classe on peut le calculer avec la formul de Sturge:  
$$K = 1 + 3.332 \log_{10}(N);$$
- ▶ l'étendu de  $X$ :  $e = x_{max} - x_{min}$ ;
- ▶ l'amplitude d'une classe:  $a_i = e_i - e_{i-1}$ , on peut le calculer avec la formule:  
$$a_i = \frac{x_{max} - x_{min}}{k} ;$$
- ▶ l'effectif d'une classe  $n_i$ : est le nombre d'individus qui prennent des valeurs  $x_i \in [e_{i-1}, e_i]$ ;



## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- ▶ Nombre de classe on peut le calculer avec la formul de Sturge:  
$$K = 1 + 3.332 \log_{10}(N);$$
- ▶ l'étendu de  $X$ :  $e = x_{max} - x_{min}$ ;
- ▶ l'amplitude d'une classe:  $a_i = e_i - e_{i-1}$ , on peut le calculer avec la formule:  
$$a_i = \frac{x_{max} - x_{min}}{k} ;$$
- ▶ l'effectif d'une classe  $n_i$ : est le nombre d'individus qui prennent des valeurs  $x_i \in [e_{i-1}, e_i]$ ;
- ▶ effectif cumulé en  $x_i$ : la somme  $N_i = \sum_{j=1}^i n_j$ ;



## Variable statistique continue

Classes	Effectifs	Fréquences absolues	Fréquences cumulées
$e_{i-1} \leq X < e_i$	$n_i$	$f_i$	$N_i$

Généralement, les données sont regroupées en  $k$  classes. Une classe est définie par ses extrémités  $e_{i-1}$ ,  $e_i$  et son effectif  $n_i$ .

- ▶ Nombre de classe on peut le calculer avec la formul de Sturge:  
$$K = 1 + 3.332 \log_{10}(N);$$
- ▶ l'étendu de  $X$ :  $e = x_{max} - x_{min}$ ;
- ▶ l'amplitude d'une classe:  $a_i = e_i - e_{i-1}$ , on peut le calculer avec la formule:  
$$a_i = \frac{x_{max} - x_{min}}{k};$$
- ▶ l'effectif d'une classe  $n_i$ : est le nombre d'individus qui prennent des valeurs  $x_i \in [e_{i-1}, e_i]$ ;
- ▶ effectif cumulé en  $x_i$ : la somme  $N_i = \sum_{j=1}^i n_j$ ;
- ▶ fréquence cumulée en  $x_i$ : la somme  $F_i = \sum_{j=1}^i f_j$ .



## Exemple 2 (Variable statistique continue)

Étude statistique de durée de vie de 100 dispositifs électronique identiques.

Durée de vie (en heures)	Nombre $n_i$ de dispositifs (fréquence absolue)	Fréquence relative $f_i$	Fréquence cumulée absolue	Fréquence cumulée relative $F_i$
$0 \leq X < 150$	30	0,30	30	0,30
$150 \leq X < 300$	15	0,15	45	0,45
$300 \leq X < 450$	12	0,12	57	0,57
$450 \leq X < 600$	10	0,10	67	0,67
$600 \leq X < 750$	8	0,08	75	0,75
$750 \leq X < 900$	8	0,08	83	0,83
$900 \leq X < 1\ 050$	8	0,08	91	0,91
$1\ 050 \leq X < 1\ 200$	6	0,06	97	0,97
$1\ 200 \leq X < 1\ 350$	3	0,03	100	1



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

Variable quantitative discrète: diagramme en bâtons

Variabes continues ou réparties en classes: Histogramme

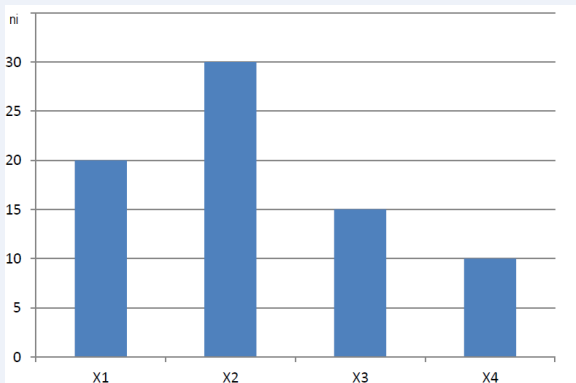
Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue



## Diagramme en tuyaux d'orgues

Pour construire un diagramme en tuyaux d'orgues, on porte sur l'axe des abscisses les modalités  $x_i$  et sur l'axe des ordonnées les effectifs  $n_i$  (ou  $f_i$ ), puis on représente les modalités par des rectangles dont la hauteur est proportionnelle à l'effectif correspondant et dont la base est une constante.



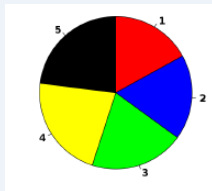


## Diagramme circulaire

Les diagrammes circulaires consistent à partager un disque, en tranches, ou secteurs, correspondant aux modalités observées. L'élaboration de ce diagramme consiste à répartir un angle de  $360^\circ$  entre les différentes modalités selon le poids (effectif) de chacune dans l'ensemble de la population. Pour obtenir le secteur correspondant à chaque modalité, on va se baser sur la règle de 3:

$$\begin{aligned} N &\longrightarrow 360^\circ \\ n_i &\longrightarrow \alpha_i \end{aligned}$$

$$\text{Donc : } \alpha_i = \frac{360n_i}{N} = 360f_i$$





## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### **Variable quantitative discrète: diagramme en bâtons**

Variabes continues ou réparties en classes: Histogramme

Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue





## Diagramme en bâtons

La représentation graphique d'une variable discrète peut s'effectuer sous la forme de diagramme en bâtons.

La valeur observée de la variable est portée sur l'axe des abscisses, et l'effectif(fréquence ) correspondante est portée sur l'axe des ordonnées. Les hauteurs des bâtons sont proportionnelles aux effectifs (fréquences).

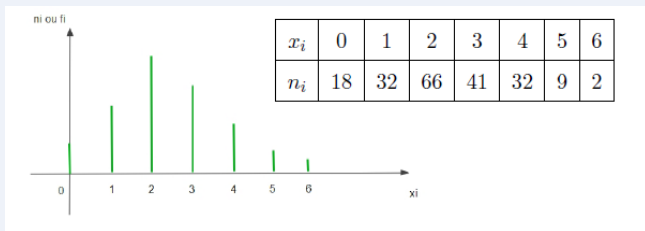


Figure: Diagramme en bâtons



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### Variable quantitative discrète: diagramme en bâtons

#### **Variabes continues ou réparties en classes: Histogramme**

#### Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue

## Histogramme

Pour chaque classe, on construit un rectangle ayant une base proportionnelle à l'intervalle de classe et dont sa surface est proportionnelles aux effectifs.

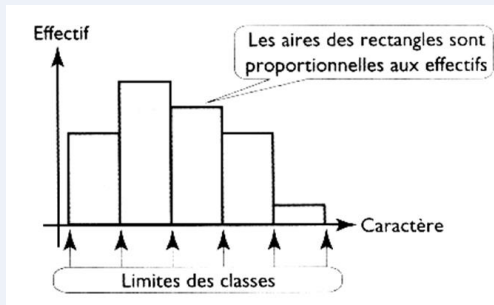


Figure: Histogramme

## Histogramme amplitudes différents

Dans le cas où les classes sont d'amplitudes différents, la hauteur du rectangle correspondant à la  $i^{\text{ème}}$  classe d'amplitude  $a_i$  sera  $h_i = n_i/a_i$ . Ainsi, la surface du rectangle représentant cette classe sera égale à  $n_i$ .

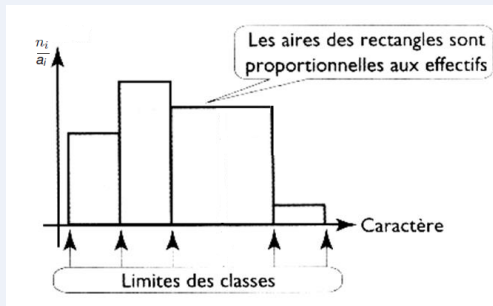


Figure: Histogramme des amplitudes différents



## Introduction

### Concepts de base de la statistique

- Population et unité statistique
- Variabes statistiques

### Rangement des données (tableau statistique)

- Variable statistique discrète
- Variable statistique continue

### Les différents modes de représentation graphique des données

#### Variable Qualitative

- Diagramme en tuyaux d'orgues(Barre)
- Diagramme circulaire

#### Variable quantitative discrète: diagramme en bâtons

#### Variabes continues ou réparties en classes: Histogramme

### Graphiques de fréquences cumulées, diagramme intégral

- Variable discrète
- Cas d'une variable statistique continue



## Variable discrète

Nous avons déjà abordé les distributions cumulées d'une variable statistique. Nous allons dans cette partie exploiter ses valeurs cumulées pour introduire la notion de la fonction de répartition.

Soit la fonction  $F_x : \mathbb{R} \rightarrow [0, 1]$  définie par:

$F_x(x) :=$  pourcentage (fréquence) des individus dont la valeur du caractère est  $\leq x$ .

Cette fonction s'appelle la fonction de répartition de la variable  $X$ . Elle satisfait, pour  $i \in \{1, \dots, n\}$ , l'expression:

$$F_x(x) = \begin{cases} 0, & \text{si } x < x_1, \\ F_1, & \text{si } x_1 \leq x < x_2, \\ \vdots & \\ F_i, & \text{si } x_i \leq x < x_{i+1}, \\ \vdots & \\ 1, & \text{si } x \geq x_n. \end{cases} \quad (1)$$



## Variable discrète

La courbe représentative de la fonction de répartition  $F_x$  s'appelle "la courbe cumulative des fréquences". elle est une courbe en escalier:

$t$	$F(t)$ (%)
$< 1$	0
$[1 ; 2[$	29,2
$[2 ; 3[$	61,0
$[3 ; 4[$	77,8
$[4 ; 5[$	92,0
$\geq 5$	100

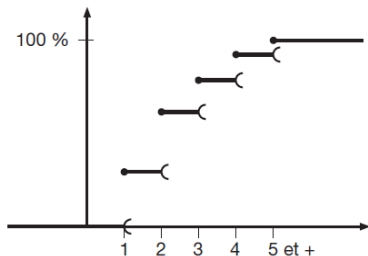


Figure: fréquence cumulé



## Variable continue

Soit la fonction  $F_x : \mathbb{R} \rightarrow [0, 1]$  définie par:

$F_x(x) :=$  pourcentage (fréquence) des individus dont la valeur du caractère est  $\leq x$ .

Cette fonction s'appelle la fonction de répartition de la variable X. Elle est donnée par l'expression (une variable continue):

$$F_x(x) = \begin{cases} 0, & \text{si } x < e_0, \\ \frac{f_1}{h}(x - e_0), & \text{si } e_0 \leq x < e_1, \\ \vdots & \\ F_i + \frac{f_{i+1}}{h}(x - e_i), & \text{si } e_i \leq x < e_{i+1}, \\ \vdots & \\ 1, & \text{si } x \geq e_n. \end{cases} \quad (2)$$





## Variable continue

La courbe de  $F_x$  est nulle avant  $e_0$ , constante égale à 1 après  $e_n$  et joint les points  $(e_0, 0)$ ,  $(e_1, F_1)$ , ...,  $(e_n, 1)$  par des segments de droites.

$t$	$F(t)$ (%)
0	0
1	16,8
3	35,8
6	50,9
12	68,7
24	87,2
36	94,2
60	100

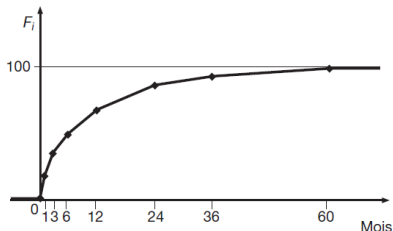


Figure: fréquence cumulé