

STATISTIQUES - GÉNÉRALITÉS

1 Vocabulaire

1.1 Population, échantillon, caractère, modalité

- On appelle univers statistique ou *population* un ensemble fini Ω dont les éléments sont des *unités statistiques* ou *individus*.
- Un sous-ensemble de Ω est appelé un *échantillon*.
- On appelle *caractère* toute application définie sur Ω à valeurs dans $\mathbb{R}$ (*caractère quantitatif*) ou dans un ensemble E dont les éléments servent à définir des situations qualitatives (*caractère qualitatif*).
- L'étude statistique porte sur un ou plusieurs caractères dont les valeurs sont les *modalités*.

1.2 Variable discrète, variable continue

- La variable est *discrète* si elle ne prend que des valeurs isolées, par exemple le nombre d'enfants par famille, nombre de pièces fabriquées par un ouvrier.
- Elle est *continue* si elle peut prendre toutes les valeurs d'un intervalle de $\mathbb{R}$.

1.3 Effectif, fréquence

- L'effectif d'une population est le nombre d'individus de cette population (on dit aussi la taille).
- Si une modalité x_i est prise par n_i individus d'une population d'effectif n , la fréquence de cette modalité est le nombre $f_i = \frac{n_i}{n}$.

2 Présentation des données

2.1 Caractère qualitatif ou quantitatif discret

Soit X un caractère qualitatif, et $\{x_1, x_2, \dots, x_k\}$ ses modalités, avec l'identité

$\sum_{i=1}^k n_i = n$, le résultat de l'étude peut se présenter sous forme :

- de données non groupées

On énumère simplement la liste des valeurs observées du caractère.

Individu numéro :	1	2	3	...	$n-1$	n
Valeurs observées	x_1	x_2	x_3		x_{n-1}	x_n

- de données groupées

Modalités	x_1	x_2	x_3	...	x_k
Nombre d'observations	n_1	n_2	n_3		n_k

2.2 Caractère quantitatif continu

Les données peuvent encore selon le même principe être non groupées, mais il s'agit de présenter un classement des données, les observations seront alors regroupées par classes de valeurs du caractère et non plus par modalités.

Classes	$[e_1, e_2[$	$[e_2, e_3[$	...	$[e_k, e_{k+1}[$
Nombre d'observations	n_1	n_2	...	n_k

2.3 Série statistique à deux caractères X et Y

Exemple 1

Effectif peu important :

Le tableau ci-dessous donne l'évolution de la consommation de médicaments des ménages en France.

- x représente l'année,
- y représente le montant de la consommation de médicaments.

x_i (Année)	1960	1965	1970	1975	1980	1985	1990	1992
y_i (milliards de francs)	3,52	6,6	10,73	20,26	33,69	64,2	95,92	108,9

Exemple 2

Tableau à double entrée

- Poids et tailles des élèves d'une classe
- X est le poids en kg
- Y est la taille en cm

$\begin{matrix} X \\ Y \end{matrix}$	$[45, 60[$	$[60, 70[$	$[70, 80[$	$[80, 100[$
$[150, 165[$				
$[165, 175[$				
$[175, 185[$				
$[185, 200[$				

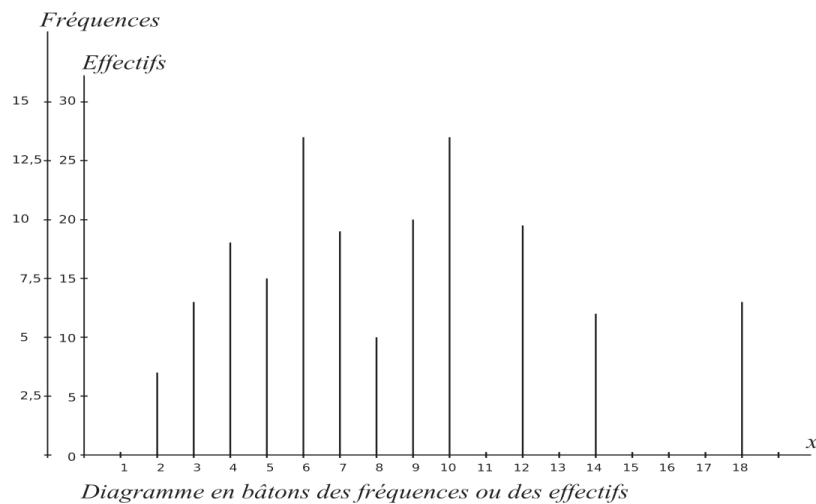
On place dans chaque case l'effectif correspondant à la classe de la ligne et à la classe de la colonne.

3 Représentations graphiques

3.1 Caractère discret

Nous cherchons à représenter la répartition des notes à une épreuve d'examen.

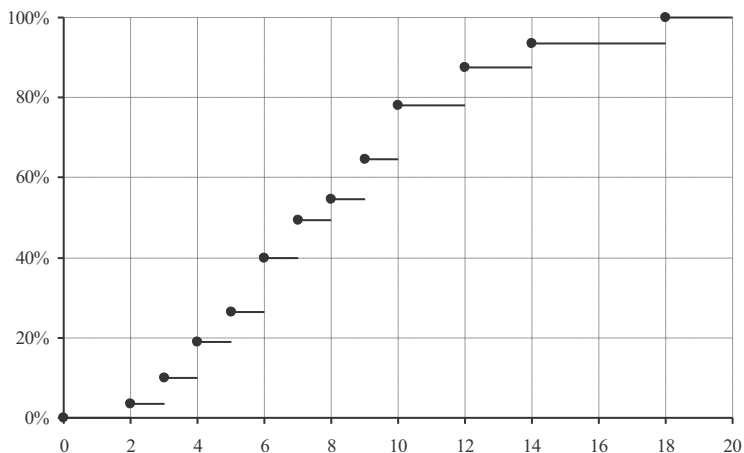
Notes x_i	2	3	4	5	6	7	8	9	10	12	14	18
Nombres d'élèves	7	13	18	15	27	19	10	20	27	19	12	13
Fréquences en %	3,5	6,5	9	7,5	13,5	9,5	5	10	13,5	9,5	6	6,5



Construction avec cumul :

Nous pouvons ajouter les lignes suivantes au tableau précédent.

Notes x_i	2	3	4	5	6	7	8	9	10	12	14	18
Effectifs cumulés croissants	7	20	38	53	80	99	109	129	156	175	187	200
Fréquences cumulées en %	3,5	10	19	26,5	40	49,5	54,5	64,5	78	87,5	93,5	100



3.2 Caractère continu, les valeurs sont groupées en classes

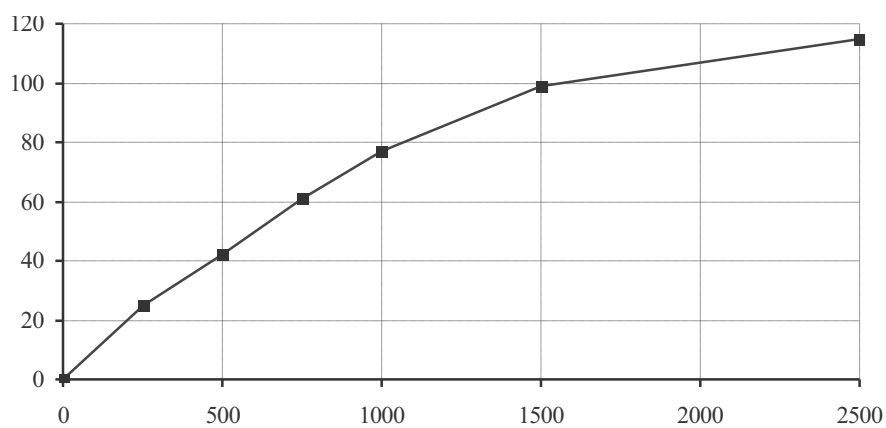
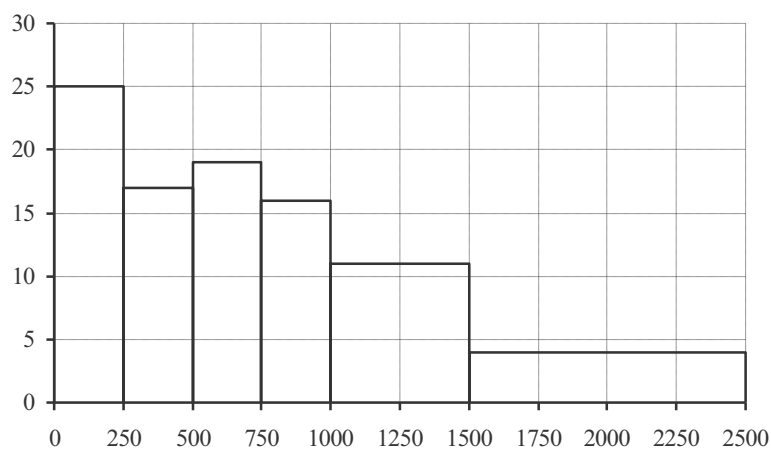
Exemple :

Un commerçant a classé les chèques encaissés au cours d'une période, suivant leur montant

Il procède à un classement selon des intervalles de valeurs.

Montant d'un chèque	Nombres de chèques	Effectifs cumulés
$[0, 250[$	25	25
$[250, 500[$	17	42
$[500, 750[$	19	61
$[750, 1000[$	16	77
$[1000, 1500[$	22	99
$[1500, 2500[$	16	115

Nous en déduisons un histogramme où la partie significative est la surface de chaque zone.



Courbe des effectifs cumulés croissants

On peut aussi utiliser les fréquences

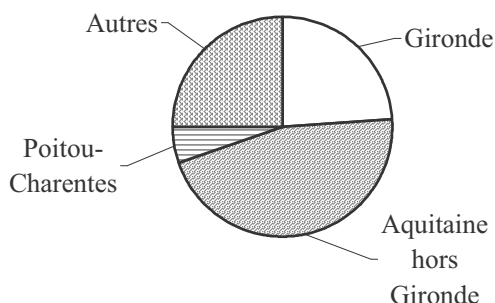
3.3 Représentation en secteurs

Sur un échantillon de 778 étudiants inscrits à un concours à Bordeaux en 1995, les effectifs suivants ont été relevés :

- 186 candidats habitent dans le département de la Gironde,
- 357 candidats habitent dans les autres départements de l'Aquitaine,
- 41 candidats habitent dans la région Poitou-Charentes,
- 194 candidats habitent hors des deux régions Aquitaine et Poitou-Charentes.

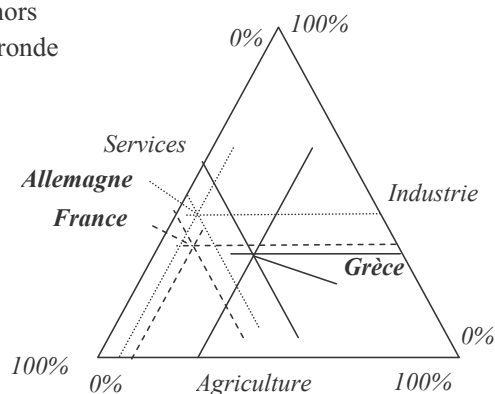
On en fait une représentation en secteurs :

- pour la Gironde, la mesure angulaire est $\frac{186}{778} \times 360 \approx 86^\circ$
- pour l'Aquitaine hors Gironde, la mesure angulaire est $\frac{357}{778} \times 360 \approx 165^\circ$
- pour la région Poitou-Charentes, la mesure angulaire est $\frac{41}{778} \times 360 \approx 19^\circ$
- pour l'extérieur des deux régions, la mesure angulaire est $\frac{194}{778} \times 360 \approx 90^\circ$



3.4 Diagramme triangulaire

Il s'agit, dans ce diagramme, d'effectuer une évaluation des trois grandes branches d'activités agriculture, industrie, services, pour en tirer des comparaisons.



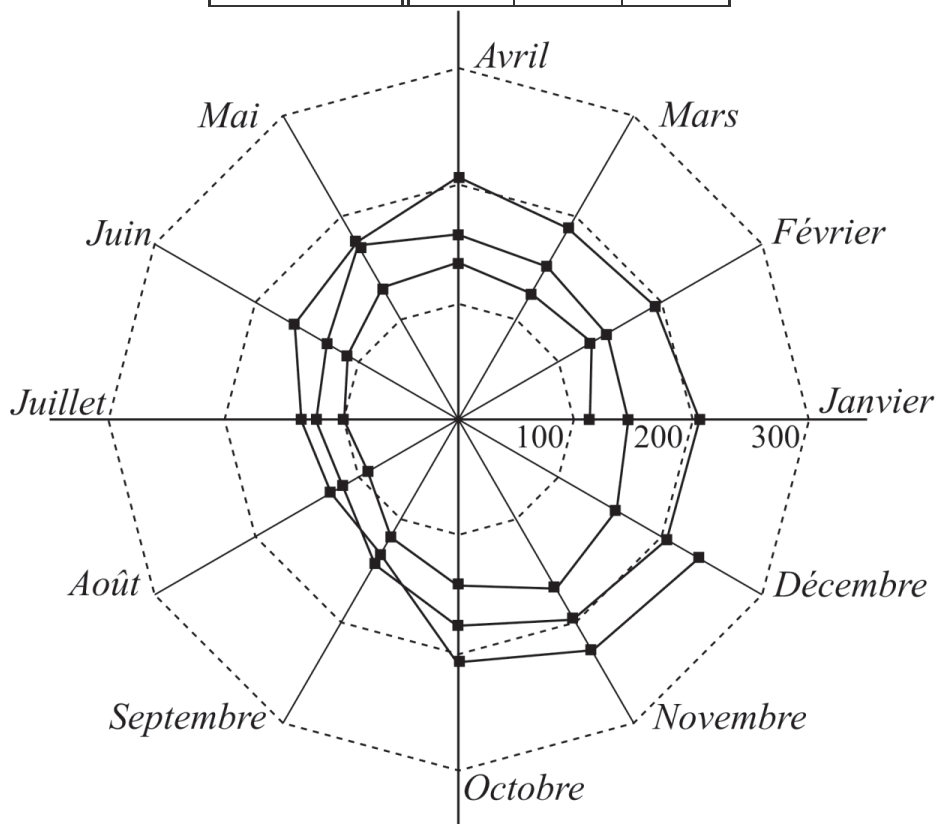
Pays	Agriculture	Industrie	Services
Grèce	30,3 %	30,2 %	39,5 %
Allemagne	6,0 %	44,8 %	49,2 %
France	8,8 %	35,9 %	55,3 %

3.5 Diagramme polaire

Ce type de diagramme est destiné aux phénomènes cycliques, annuels, hebdomadaires...

Dans cet exemple, nous étudions les indices de consommation d'électricité dans une entreprise.

Mois	1998	1999	2000
Janvier	112	145	206
Février	130	145	195
Mars	122	150	187
Avril	134	158	205
Mai	130	173	175
Juin	110	130	163
Juillet	98	123	135
Août	88	115	125
Septembre	115	140	135
Octobre	140	175	205
Novembre	165	195	225
Décembre	155	205	235



EXERCICES

Exercice 1.1

Les notes de 1 000 étudiants lors d'une épreuve de statistique se répartissent de la manière suivante :

Notes x_i	6	7	8	9	10	11	12	13	14	15	16	17	18
Effectifs n_i	10	23	45	78	116	147	162	148	117	77	46	20	11

1°) Présenter un tableau où figureront entre autres :

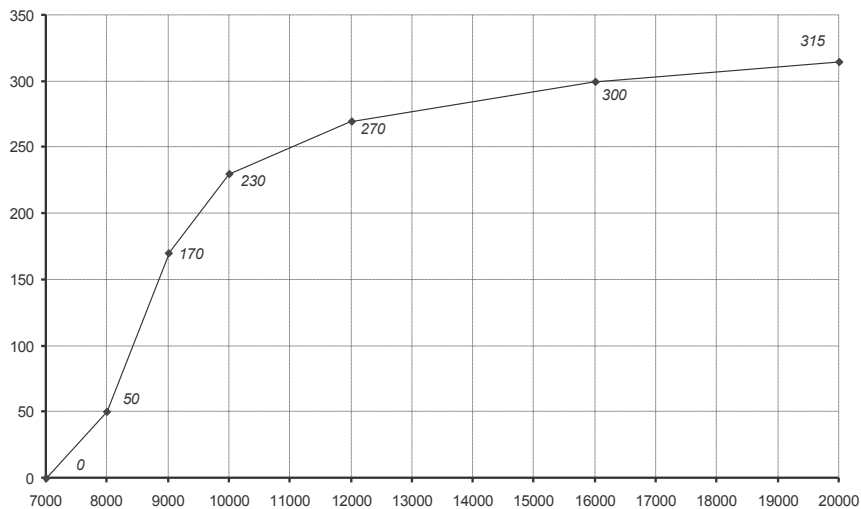
- les fréquences
- les effectifs cumulés croissants
- les fréquences cumulées croissantes.

2°) Effectuer un diagramme en bâtons des fréquences de la série statistique.

3°) Tracer la courbe des fréquences cumulées croissantes.

Exercice 1.2

Les salaires nets des employés d'une entreprise ont permis



d'établir la courbe des effectifs cumulés croissants

1°) Présenter le tableau où figureront entre autres

- les classes
- les effectifs
- les effectifs cumulés croissants
- les effectifs cumulés décroissants

2°) Présenter l'histogramme de cette série

Exercice 1.3

Durant une période déterminée, les montants des dépôts sur les livrets d'épargne dans une banque se répartissent comme suit :

- 1°) Représenter graphiquement cette série.
- 2°) Combien de dépôts inférieurs à 1200 francs ont été effectués durant cette période?
- 3°) Tracer la courbe des effectifs cumulés croissants et vérifier le calcul des quartiles de l'exercice 2.3

Montant des dépôts en francs	Nombre de dépôts
[0,400[	123
[400,600[	217
[600,800[	148
[800,1000[	156
[1000,1200[	226
[1200,1600[	188
[1600,2000[	122
[2000,3000[	102

Exercice 1.4

On dispose du poids en kilogramme d'un échantillon de 200 bêtes d'une variété d'animaux adultes.

- 1°) Représenter graphiquement cette série.
- 2°) Tracer la courbe des effectifs cumulés croissants et vérifier le calcul des quartiles de l'exercice n°2.4.

Classes en kg	Effectifs
[55,65[	10
[65,75[	23
[75,77[	21
[77,79[	28
[79,81[	34
[81,83[	28
[83,85[	25
[85,95[	23
[95,105[	8

Exercice 1.5

La répartition des salaires des employés d'une entreprise est donnée dans le tableau suivant :

Salaires en francs	Centres	Effectifs	Fréquences(%)	Effectifs cumulés croissants
[6000 ; 7600 [			5	
[7600 ; 8400 [			8,5	
[8400 ; 9200 [			23,75	
[9200 ; [			25	
[; [	10800			
[11600 ; 14000 [			6,75	400

- 1°) Compléter le tableau.
- 2°) Combien d'employés ont un salaire inférieur à 9200F?
- 3°) Effectuer un histogramme des effectifs de la série statistique.