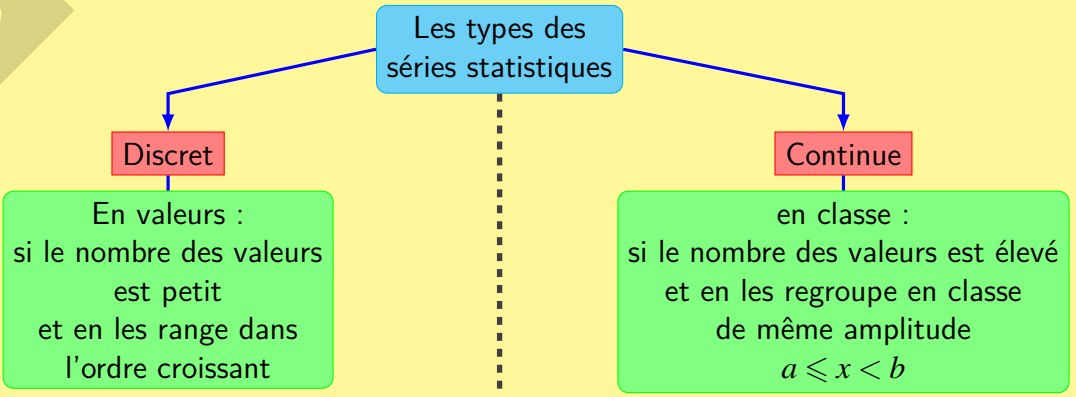
Statistiques

I Rappel

① Étude statistique : étude d'un phénomène ou d'un caractère qui caractèrise les membres d'un groupe
② Population statistique : c'est l'échantillon ou le groupe qui fait l'objet de l'étude statistique et chaque membre s'appelle individu ou unité statistique
③ Caractère : c'est le phénomène étudié qui est une propriété qu'on peut observer ou mesurer et on a deux types de caractère
★ Caractère quantitatif : caractère qu'on peut exprimer par des nombres (nombres d'enfants, les notes, âge, pages, longueurs,...)
★ Caractère qualitatif : caractère qu'on ne peut pas exprimer par des nombres (sex, couleur, type de voiture,...)
④ Effectif : c'est le nombres des unités qui prennent une des valeurs du caractère, on le symbolise par n_i (nombres de répétition)
⑤ Effectif total : c'est la somme de tous les effectifs symbolisé par N
⑥ Effectif cumulé : l'eefectif cumulé d'une valeur du caractère est la somme des effectifs de toutes les valeurs qui sont inférieur ou égale à cette valeur
⑦ Fréquence : la fréquence d'une valeur est le quotient de l'effectif de cette valeur par l'effectif total, on la note : $f_i = \dfrac{n_i}{N}$
⑧ Fréquence cumulée : la fréquence cumulée d'une valeur est le quotient de l'effectif cumulé de catte valeur par l'effectif total N
⑨ Pourcentage : c'est le produit de la fréquence par 100
on le note : $p_i = \dfrac{\text{Effectifs cumulé}}{\text{Effectif total}} \times 100 = f_i \times 100$

★ La somme de toutes les fréquences d'une série statistique est égale à 1
★ La fréquence cumulée d'une valeur est la somme de la fréquence de cette valeur et des fréquences de toutes les valeurs précédentes
★ On a deux types de séries statistiques :



Yassine Elamri

II Effectifs cumulés, fréquences et fréquence cumulées

1 Série statistique discrète en valeur

Si le caractère est quantitatif et le nombre de ses valeurs est petit, on les range dans l'ordre croissant

Application

Le tableau, ci-après, représente une série statistique qui exprime la répartition de 24 adhérents dans un club sportif selon leur âge

Caractère	12	13	14	15	16
Effectifs	5	6	x	8	4
Effectifs cumulés	...	11	...	...	24
Fréquences	...	...	...	...	...
Fréquences cumulées	...	...	...	...	...

1. Quelle est la population statistique de cette série statistique ?
2. Quelle est l'unitè statistique de cette série statistique ?
3. Quel est le caractère étudié dans cette série statistique ? quel est son type ?
4. Déterminer la valeur de x
5. Compléter le tableau
6. Vérifier que la sommes des fréquences est égale à 1
7. Représenter cette série statistique avec un diagramme en bâtons et un diagramme à ligne brisée

Solution

1. La population statistique de cette série statistique est 24 adhérent dans un club sportif
2. L'unitè statistique de cette série statistique est un adhérent
3. Le caractère étudié dans cette série statistique est l'âge de l'adhérent qui est un caractère quantitatif discret
4. Déterminons x
   On a l'effectifs total est $N = 24$,c'est à dire $5+6+x+8+4 = 24$, donc $23+x = 24$
   Alors $x = 24-23 = 1$

Yassine Elamri

Today better than yesterday, tomorrow better than today

5 Complétons le tableau

Caractère	12	13	14	15	16
Effectifs	5	6	1	8	4
Effectifs cumulés	5	11	12	20	24
Fréquences	$\frac{5}{24}=0.21$	$\frac{6}{24}=0.25$	$\frac{1}{24}=0.04$	$\frac{8}{24}=0.33$	$\frac{4}{24}=0.17$
Fréquences cumulées	0.21	0.46	0.50	0.83	1

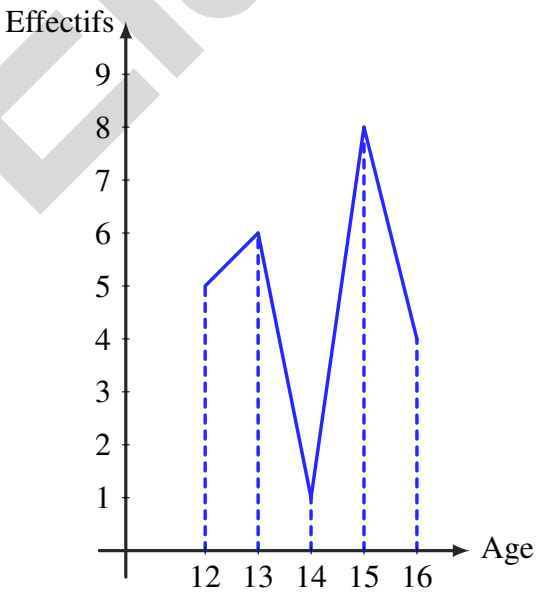
6 Vérifions la sommes des fréquences
On a $0.21+0.25+0.04+0.33+0.17=1$, alors la somme de toutes les fréquences est égale à 1

7 Représentation de la série statistique

✪ Avec un diagramme en bâtons



✪ Avec un diagramme à ligne brisée (courbe polygonale)



Application

Une étude a été menée sur le nombre d'enfants dans 20 familles, et a donnée les résultats suivants : $2-3-4-3-0-4-3-2-1-1-2-1-0-2-3-4-1-3-0-1$

1. Donner le tableau des effectifs et des effectifs cumulés de cette série statistique
2. Calculer la fréquence de la valeur 0
3. Calculer le pourcentage de la valeur 0
4. Calculer le pourcentage du nombre de familles dont le nombre d'enfants dépasse 2
5. Représenter cette série en collones

Yassine Elamri

Today better than yesterday, tomorrow better than today

Solution

1 Le tableau des effectifs et des effectifs cumulés de la série

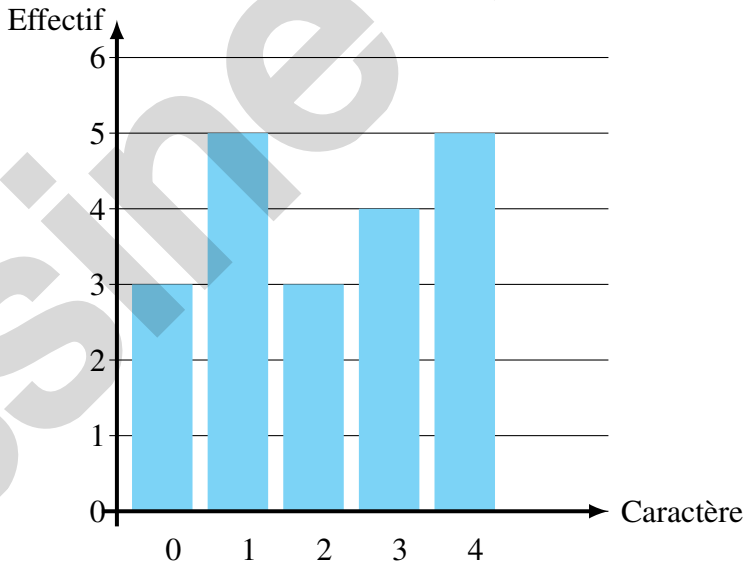
Caractère	0	1	2	3	4
Effectifs	3	5	4	5	3
Effectifs cumulés	3	8	12	17	20

2 Calculeons la fréquence de la valeur 0
Soit f la fréquence de la valeur 0
Donc $f = \frac{n}{N} = \frac{3}{20} = 0.15$

3 Calculons le pourcentage de la valeur 0
Soit p le pourcentage de la valeur 0
Donc $p = f \times 100 = 0.15 \times 100 = 15\%$

4 Calculons le pourcentage des familles dont le nombre d'enfants dépasse 2
Le nombres de familles qui ont plus de deux enfants est $5+3=8$
Donc le pourcentage est : $P = \frac{8}{20} \times 100 = \frac{800}{20} = 40\%$

5 Représentons cette série en collones



2 Série statistique en classes

Si le caractère est quantitatif et le nombre de ses valeurs est élevé, au lieu d'étudiertoutes les valeurs, on les regroupe dans des intervalles de même amplitude $[a,b[$ appelés des classes et le centre de la classe est $\frac{a+b}{2}$

Yassine Elamri

Today better than yesterday, tomorrow better than today

Application

Le bilan suivant donne la répartition d'âges des ouvriers dans une ferme :
16 – 26 – 34 – 17 – 22 – 45 – 36 – 27 – 29 – 25 – 19 – 18 – 32 – 42 – 21 – 33 – 35 – 16 – 26
34 – 17 – 22 – 38 – 36 – 27 – 29 – 38 – 13 – 18 – 32 – 30 – 39

1. Déterminer la population de cette série
2. Déterminer le caractère étudié et son type
3. Compléter le tableau suivant

Age en années	$[10,20[$	$[20,30[$	$[30,40[$	$[40,50[$
Centre de la classe	...	...	...	...
Effectifs : Nombre d'ouvriers	...	...	...	...

4. Combien y-a-t-il d'ouvriers dans la ferme ?
5. Calculer le pourcentage des ouvriers dont l'âge est inférieur à 20 ans
6. Calculer la fréquence de la classe $[30,40[$
7. construire l'histogramme de la répartition des ouvriers de la ferme selon les classes de leurs âges

Solution

1. La population de cette série est les ouvriers de la ferme
2. Le caractère statistique étudié est l'âge des ouvriers et c'est un caractère quantitatif continu
3. Complétons le tableau

Age en années	$[10,20[$	$[20,30[$	$[30,40[$	$[40,50[$
Le centre de la classe	$\frac{10+20}{2}=15$	$\frac{20+30}{2}=25$	35	45
Effectifs : Nombre d'ouvriers	8	10	12	2

4. Le nombre d'ouvriers dans la ferme est : $N = 8+10+12+2 = 32$
5. Le pourcentage des ouvriers dont l'âge est inférieur à 20 ans
   Le nombre d'ouvriers qui ont une âge inférieur à 20 ans est $n = 8$
   Donc leur pourcentage est : $p = \frac{n}{N} \times 100 = \frac{8}{32} \times 100 = 25\%$
6. La fréquence de la classe $[30,40[$ est $f = \frac{n}{N} = \frac{12}{32} = 0.375$
7. construisons l'histogramme de la répartition
   Si la répartition est sous forme de classes $[a,b[$ ou d'intervalles $a \leqslant x < b$, on emploie un histogramme

Yassine Elamri

Today better than yesterday, tomorrow better than today

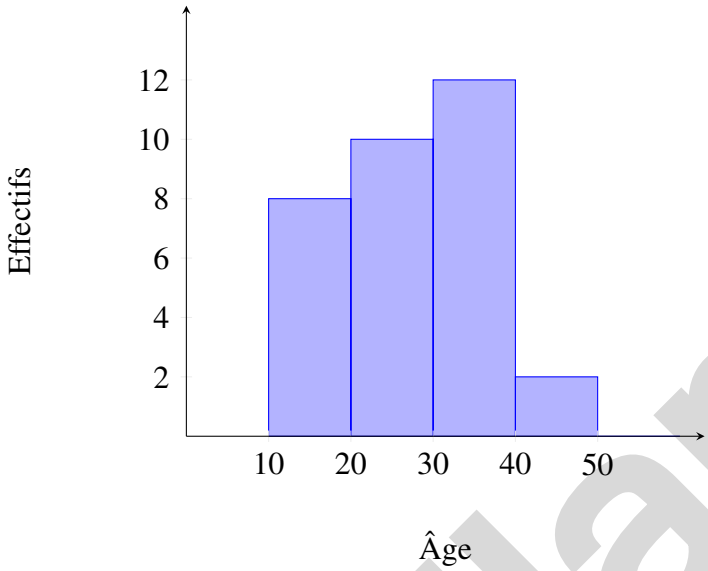


Diagramme circulaire

Ce type de diagramme est concenable pour les caractères qualitatifs.
Un diagramme circulaire est un diagramme qui a la forme d'un disque (ou un demi-disque), décomposé en secteurs dont les mesures des angles sont proportionnelles aux efectifs (et également aux fréquences)

$$\begin{aligned}\text{Mesure de l'angle} &= \frac{\text{Effecifts}}{\text{Effecift total}} \times 360^\circ \\ &= \text{Fréquence} \times 360^\circ \\ &= \frac{\text{Pourcentage}}{100} \times 360^\circ\end{aligned}$$

III Les paramètres de position

1 Le mode

a Série en valeurs :

Définition

Le mode d'une sèrie statistique est la valeur du caractère qui a le plus grand effectif

• Exemple

➪ Exemple ❶

Yassine Flamri

Today better than yesterday, tomorrow better than today

Caractère	12	13	14	15	16
Effectifs	5	6	1	8	4

On a le plue grand effectif est : 8, sa valeur est 15
Donc le mode de cette sèrie est 15 ⇨ Exemple ❷

Caractère (coefficient)	1	2	3	5
Effectifs (Nombre de matières)	3	2	2	3

Pour cette sèrie, le plue grand effectif est : 3 correspendant aux valeurs 1 et 5
Donc cette sèrie admet deux modes : 1 et 5
⇨ Exemple ❸
Les longueurs de 8 élèves en 'cm' sont : 61;68;67;73;66;64;59;70
Existe-t-il une valeur plus fréquente que les autres dens ces données ?
La réponse est *non*
Donc, cette sèrie n'admet pas de mode

b Série en classes :

Définition

On appelle classe modale, d'une série statistique groupée en classe, toute classe qui a le plus grand effectif

• Exemple

On considère la série suivante

Classes	[120;130[	[130;140[	[140;150[	[150;160[
Effectifs	9	11	12	18

Le plus grand effectif est 18, et il correspond à la classe [150;160[, donc la classe modale est : [150;160[

Remarque

Une série statistique peut ne pas avoir un mode (classe modale), comme elle peut avoir plusieurs modes (classes modales)

Yassine Elamri

Today better than yesterday, tomorrow better than today

2 La moyenne arithmétique

a Série en valeurs

Définition

La moyenne arithmétique est le rapport de la somme de tous les produit de chaque valeur (centre de classe) par son effectif sur l'effectif total, on la note m

Remarque

La moyenne arithmétique m est égale au rapport de la somme de toutes les valeurs du caractère sur l'effectif total, c'est à dire, c'est la valeur obtenue si toutes les valeurs du caractère sont égaux

Exemple

❶ Regarder l'application 2

On a $m = \dfrac{(0\times 3)+(1\times 5)+(2\times 4)+(3\times 5)+(4\times 3)}{20} = \dfrac{0+5+8+15+12}{20} = \dfrac{40}{20} = 2$

Donc 2 est la moyenne du nombres d'enfants dans chaque famille

❷ On considère la série suivante

Caractère (Coefficient)	1	2	3	5
Effectifs (Nombres de matières)	3	2	2	3

On a $m = \dfrac{(1\times 3)+(2\times 2)+(3\times 2)+(5\times 3)+}{10} = \dfrac{3+4+6+15}{10} = \dfrac{28}{10} = 2.8$

Donc 2.8 est la moyenne de cette série

b Série en classes

Règle

Si $a \leqslant x < b$ est une classe d'une sérire, alors son milieu est $\dfrac{a+b}{2}$

Pour calculer la moyenne d'une série statistique en classes, on utilise la définition précédente en remplaçant les valeurs par les centres des classes

Exemple

↦ Regarder l'application 3

Age en années	$[10,20[$	$[20,30[$	$[30,40[$	$[40,50[$
Le centre de la classe	15	25	35	45
Effectifs : Nombre d'ouvriers	8	10	12	2

Yassine Elamri

Today better than yesterday, tomorrow better than today

$$m = \frac{(15\times 8)+(25\times 10)+(35\times 12)+(45\times 2)}{32} = \frac{120+250+420+90}{32} = 27.5$$

Alors, la moyenne d'âges des ouvriers de la ferme est 27.5 et cela signifie que si on a supposé que tous les ouvriers ont le même âge, l'âge de chacun d'eux va être 27.5 ans

3 La médiane

Définition

La médiane d'une série statistique est la plus petite valeur du caractère dont l'effectif cumulé est supérieur ou égal à la moitié de l'effectif total

• Exemple

✯ série en valeurs :

Regarder l'application 1

Caractère	12	13	14	15	16
Effectifs	5	6	1	8	4
Effectifs cumulés	5	11	12	20	24

La moitié de l'effectif total est $\frac{24}{2} = 12$

La plus petite valeur dont l'effectif cumulé est supérieur ou égal à 12 est 14

Donc, la médiane de cette série est égale à 14 **✯ série en classes :**

Classes	$[120;130[$	$[130;140[$	$[140;150[$	$[150;160[$
Effectifs	9	11	12	18
Effectifs cumulés	9	20	32	50

La moitié de l'effectif total est $\frac{50}{2} = 25$

La premier effectif cumulé supérieur ou égal à 25 est 32, qui correspond à la classe $[140;150[$

Donc, la médiane de cette série se trouve dans la classe $[140;150[$

Remarque

On peut dire (admettre) que 145 (centre de la classe $[140;150[$) est la médiane de cette série statistique

Définition (Autre définition)

La médiane d'une série statistique, dont les valeurs sont ordonnées est la valeur du caractère qui partage la série en deux parties de même effectif

Yassine Elamri

Today better than yesterday, tomorrow better than today

• Exemple

★ **Cas ①** : Effectifs total N impair

Le nombre d'abssence pendant 7 jours des ouvriers d'une sociètè sont : $3-1-2-0-4-2-3$

On les range dans l'ordre croissant : $0-1-2-2-3-3-4$

$\underbrace{0-1-2}_{3\text{ valeurs}}-\underset{\uparrow}{\underset{M}{2}}-\underbrace{3-3-4}_{3\text{ valeurs}}$

Donc, la médiane est 2

★ **Cas ②** : Effectif total N pair

Le nombre d'abssence pendant 8 jours des ouvriers d'une sociètè sont : $0-1-1-2-3-4-4-5$

On les range dans l'ordre croissant : $0-1-1-2-3-4-4-5$

$\underbrace{0-1-1-2}_{4\text{ valeurs}}-\underset{\uparrow}{\underset{M=\frac{2+3}{2}}{}}-\underbrace{3-4-4-5}_{4\text{ valeurs}}$

On peut prendre comme médiane tout nombre entre 2 et 3, d'ou, la médiane de cette série est 2.5

IV La dispersion

Définition

On considère deux séries statistiques S_1 et S_2, qui ont la même moyenne arithmétique m

On dit que la série S_1 est moins dispersée que la série S_2 si les valeurs de la série S_1 sont plus proches de m que celles de la série S_2

• Exemple

On considère le tableau suivant :

Devoirs	D_1	D_2	D_3	D_4	D_5
Notes de Hassan	9	14	10	13	14
Notes de Khalid	8	15	10	17	9

La moyenne de *Hassan* est : $m_1 = \dfrac{9+14+10+13+14}{5} = \dfrac{60}{5} = 12$

La moyenne de *Khalid* est : $m_2 = \dfrac{8+16+10+17+9}{5} = \dfrac{60}{5} = 12$

Donc $m_1 = m_2$, c'est à dire que Hassan et Khalid ont la même moyenne 12

On remarque que les notes de Hassan sont plus proches de la moyenne 12 que celles de Khalid

On dit que les notes de Hassan sont moins dispersées que les notes de Khalid

Yassine Elamri

Today better than yesterday, tomorrow better than today