MINISTÈRE DE L'EDUCATION NATIONALE
ET DE L'ALPHABETISATION

REPUBLIQUE DE COTE D'IVOIRE

Union – Discipline – Travail

MON ECOLE A LA MAISON

SECONDAIRE

1ère D
MATHEMATIQUES

CÔTE D'IVOIRE – ÉCOLE NUMÉRIQUE

Durée : 07 heures Code :

Compétence 2 **Traiter des situations relatives à la modélisation de phénomènes aléatoires, à l'organisation et aux traitements de données**

Thème 1 **Organisation et traitements de données**

Leçon 15 : STATISTIQUES

A. SITUATION D'APPRENTISSAGE

L'équipe de course à pied d'un lycée a un nouvel entraîneur. Celui-ci vient de recevoir le tableau ci-dessous indiquant le temps mis par chacun des membres de l'équipe lors de la dernière épreuve de 10 km.

Nom	Temps (en min)
Agnero	53
Aka	51
Akalé	66
Allou	63
Amani	59
Ballo	61
Camara	48
Dago	41
Ehouman	47
Fallé	46

Nom	Temps (en min)
Goly	51
Gnali	60
Kassi	49
Koffi	46
Kouamé	44
Kouman	43
Lath	52
Lamine	39
Lohess	42
Manouan	53

Nom	Temps (en min)
Pakora	51
Sery	57
Seyo	62
Tiékoura	50
Traoré	43
Vanié	47
Yao	48
Yéo	56
Zadi	49
Zatto	61

Soucieux d'améliorer les performances de l'équipe, l'entraîneur expose ses décisions suivantes à l'équipe. « Je vais vous partager en cinq équipes de même effectif et de niveau équivalent (selon le temps mis lors de votre dernière épreuve).
Pour exposer les raisons de mon choix, je vais faire un affichage présentant une représentation graphique sous forme d'un histogramme.
Chacun des sportifs sera situé par rapport aux autres avec le classement, ainsi qu'une mise en évidence du premier quart, de la moitié et du troisième quart et des temps correspondants ».
Les élèves des classes de première scientifique faisant partie de l'équipe sont impatients de savoir dans quelles équipes ils seront et quelle est la situation de chacun par rapport aux autres.
Ils se mettent ensemble à organiser les données ci-dessus pour répondre à leurs préoccupations.

B. RESUME DE COURS

I. SERIES STATISTIQUES REGROUPEES EN CLASSES

1. Rappel

Lorsqu'il est question d'une série statistique regroupée en classes, on considère le tableau suivant :

Valeurs de X	$[x_1; x_2[$	$[x_2; x_3[$	$[x_3; x_4[$	…	$[x_p; x_{p+1}[$	TOTAL
Effectifs	n_1	n_2	n_3	…	n_p	N
Centre	c_1	c_2	c_3	…	c_p	

- L'amplitude de la classe $[x_i; x_{i+1}[$ est $x_{i+1} - x_i$
- Le centre $c_i = \frac{x_i + x_{i+1}}{2}$

Exemple :

On a relevé dans une agence bancaire les montants en milliers de francs des 49 premiers versements effectués au guichet. On a obtenu les résultats suivants :

950	300	100	800	200	30	75
250	600	260	150	45	490	400
200	375	360	620	130	1450	880
1025	560	350	450	400	280	190
1180	220	520	120	900	110	350
600	850	290	1400	125	900	1000
130	1100	430	950	45	310	590

Regroupons ces montants par classes d'amplitude 300, la première étant [0 ;300[et dressons le tableau des effectifs comportant les centres.

Classes	[0 ;300[	[300 ; 600[	[600 ; 900[	[900 ; 1200[	[1200 ;1500[
Effectifs	19	14	6	8	2
Centres	150	450	750	1050	1350

2. Densité

Définition

On appelle densité d'une classe, le quotient de l'effectif de la classe par l'amplitude de cette classe.

Exemple

Cette série statistique représente la production en tonnes de plusieurs coopératives de planteurs de cacao.

Classes	[0 ;30 [	[30 ; 45[	[45 ; 70[	[70 ; 80[	[80 ; 100[
Effectifs	13	25	20	15	17

Déterminons l'amplitude et la densité de chaque classe.

Classes	[0 ;30 [	[30 ; 45[	[45 ; 70[	[70 ; 80[	[80 ; 100[	Total
Effectifs	13	20	25	15	17	90
Amplitude	30	15	25	10	20	
Densité	0,43	1,33	1	1,5	0,85	

La densité de la classe [0 ;30 [est $\frac{13}{30} = 0{,}43$

II. Caractéristiques de position d'une série statistique regroupée en classes

1. Classe modale

Définition

On appelle classe modale toute classe dont la densité est maximale

Exemple

Cette série statistique représente la production en tonnes de plusieurs coopératives de planteurs de cacao.

Classes	[0 ;30 [	[30 ; 45[	[45 ; 70[	[70 ; 80[	[80 ; 100[
Effectifs	13	20	25	15	17
Densité	0,43	1,33	1	1,5	0,85

La classe modale de cette série statistique est [70 ; 80[

Interprétation : La classe [70 ; 80[comporte la plus grande concentration de coopératives.

Cas particulier :

Si toutes les classes ont la même amplitude, alors une classe modale est une classe dont l'effectif est maximal.

2. Moyenne

La moyenne d'une série statistique, notée : $\overline{x}$, est donnée par :

$$\overline{x} = \frac{\sum_{i=1}^{p} n_i c_i}{N} = \frac{1}{N}(n_1 c_1 + n_2 c_2 + \cdots + n_p c_p)\ .$$

En utilisant les fréquences, on a :

$$\overline{x} = \sum_{i=1}^{p} f_i c_i = f_1 c_1 + f_2 c_2 + \cdots + f_p c_p.$$

Exercice de fixation

Cette série statistique représente la production en tonnes de plusieurs coopératives de planteurs de cacao.

Classes	[0 ;30 [	[30 ; 45[	[45 ; 70[	[70 ; 80[	[80 ; 100[
Effectifs	13	20	25	15	17

Calcule la moyenne de la série statistique.

Solution

Classes	[0 ;30 [	[30 ; 45[	[45 ; 70[	[70 ; 80[	[80 ; 100[	Total
Effectifs	13	20	25	15	17	90
Centre	15	37,5	57,5	75	90	
$n_i c_i$	195	750	1437,5	1125	1530	5037,5

$n_1 c_1 = 13 \times 15 = 195$

La moyenne est : $\bar{x} = \frac{5037,5}{90} = 55,97$

Interprétation

La production moyenne de l'ensemble des coopératives est 55,97 tonnes.

3. Médiane

Définition :

La médiane d'une série statistique continue est un nombre qui sépare les valeurs ordonnées de la série en deux familles de même effectif.
Autrement dit : c'est un nombre M tel qu'au moins 50% des individus aient une valeur du caractère supérieure ou égale à M.

Elle se détermine :

- Soit graphiquement : c'est l'abscisse du point de la courbe cumulative des effectifs (resp. fréquences) dont l'ordonnée est la moitié de l'effectif total (resp. 0,5) ;

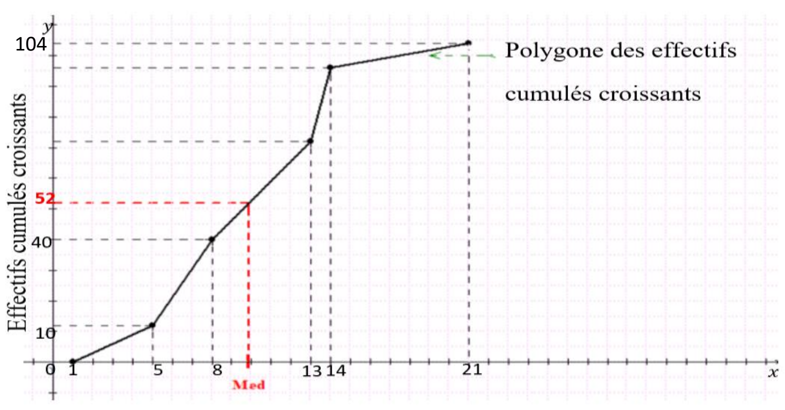
Exemple :

On a relevé le nombre d'heures d'absences de tous les élèves de $1^{ère}$ C d'un lycée durant l'année scolaire.

On a obtenu la série statistique suivante :

Classes	[1 ; 5[	[5 ; 8[	[8 ; 13[	[13 ; 14[	[14 ; 21[
Effectifs	12	28	32	24	8
Effectifs cumulés croissants	12	40	72	96	104

Le polygone des effectifs cumulés croissants est le suivant :



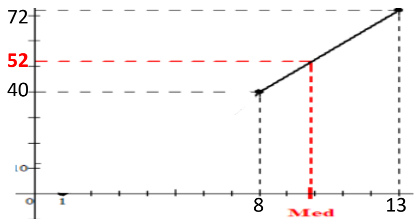
La médiane est environ égale à 9.

- La détermination de la médiane par calcul se fait par interpolation linéaire.

Exemple :

Considérons l'exemple précédent.

On a : $\frac{N}{2} = \frac{104}{2} = 52$; 52 est compris entre le 40^{e} rang et le 72^{e} rang. On a Me $\in$ [8 ; 13[.



On obtient le tableau suivant :

Modalités (Abscisse)	8	Me	13
Ecc (Ordonnée)	40	52	72

Alors

$\frac{Me-8}{13-8} = \frac{52-40}{72-40}$, donc Me = 9,87

Interprétation

La moitié des élèves a un nombre d'heures d'absences supérieur ou égal à 9,87.

Remarque : La médiane peut être déterminée graphiquement à l'aide du polygone des effectifs cumulés décroissants.

4. Quartiles

Définition :

Les valeurs de la série étant ordonnées :

- Le premier quartile, noté Q_1, est la valeur de la variable telle que 25% des valeurs sont inférieures ou égales à Q_1 et 75% lui sont supérieures.
- Le deuxième quartile Q_2, est la médiane.
- Le troisième quartile, noté Q_3, est la valeur de la variable telle que 75% des valeurs sont inférieures ou égales à Q_3 et 25% lui sont supérieures.

➢ Graphiquement,

- Q_1 correspond à 25% de l'effectif sur le polygone des effectifs cumulés croissants ou des fréquences cumulées croissantes.
- Q_3 correspond à 75% de l'effectif sur le polygone des effectifs cumulés croissants ou des fréquences cumulées croissantes.

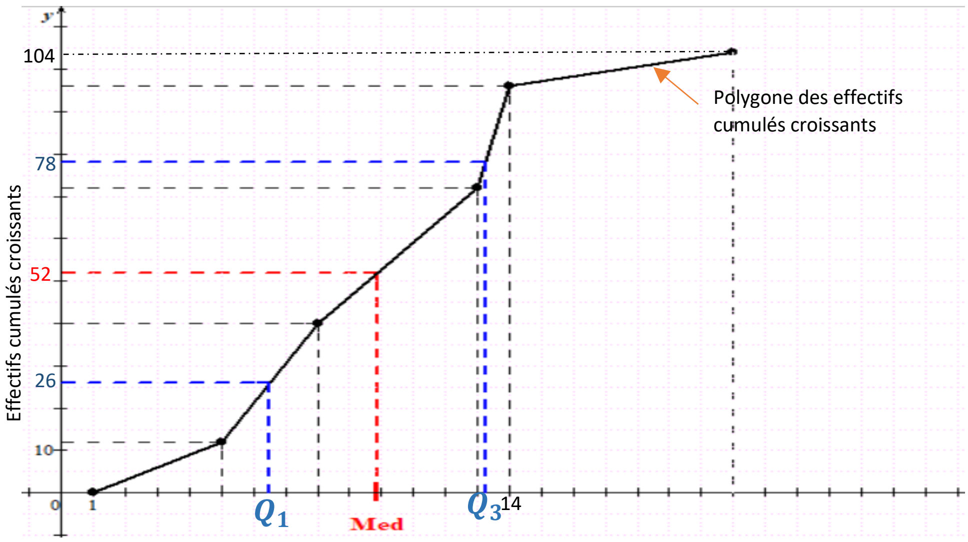
Exemple :

On a relevé le nombre d'heures d'absences de tous les élèves de 1ereC d'un lycée durant l'année scolaire.

On a obtenu la série statistique suivante .

Classes	[1 ; 5[	[5 ; 8[	[8 ; 13[	[13 ; 14[	[14 ; 21[
Effectifs	12	28	32	24	8
Effectifs cumulés croissants	12	40	72	96	104

Le polygone des effectifs cumulés croissants est le suivant :



Le premier quartile est environ égal à 7.

Le troisième quartile est environ égal à 13.

- On peut calculer une valeur plus précise de Q_1 et Q_3 par interpolation linéaire.

Calculons le premier quartile

N$\times \frac{25}{100} = 104 \times \frac{25}{100} = 26$; 26 est compris entre le 12e rang et le 40e rang. On a $Q_1 \in [5; 8[$

Alors $\frac{Q_{1-5}}{8-5} = \frac{26-12}{40-12}$, donc $Q_1 = 6{,}5$

Calculons le troisième quartile

N$\times \frac{75}{100} = 104 \times \frac{75}{100} = 78$; 78 est compris entre le 72e rang et le 96e rang. On a $Q_3 \in [13; 14[$

Alors $\frac{Q_{3-13}}{14-13} = \frac{78-72}{96-72}$, donc $Q_1 = 13{,}25$

Interprétation

- 26 élèves ont un nombre d'heures d'absences inférieur ou égal à 6,5.
- 78 élèves ont un nombre d'heures d'absences inférieur ou égal à 13,5

x

III. REPRESENTATIONS GRAPHIQUES

1. Histogramme

- L'histogramme d'une série statistique regroupée en classes est constitué de rectangles juxtaposés.
- L'aire de chaque rectangle est proportionnelle à l'effectif (resp. la fréquence) de la classe correspondante.
- Les largeurs des rectangles sont proportionnelles aux amplitudes des classes.
- Les hauteurs des rectangles sont proportionnelles aux densités des classes

Exercice de fixation

On donne la série statistique suivante :

Modalité	[2 ; 3[	[3 ; 4,5[	[4,5 ; 5,5[	[5,5 ; 6[	[6 ; 8[
Fréquence	0,15	0,35	0,25	0,1	0,15
Amplitude					
Centre					
Densité					

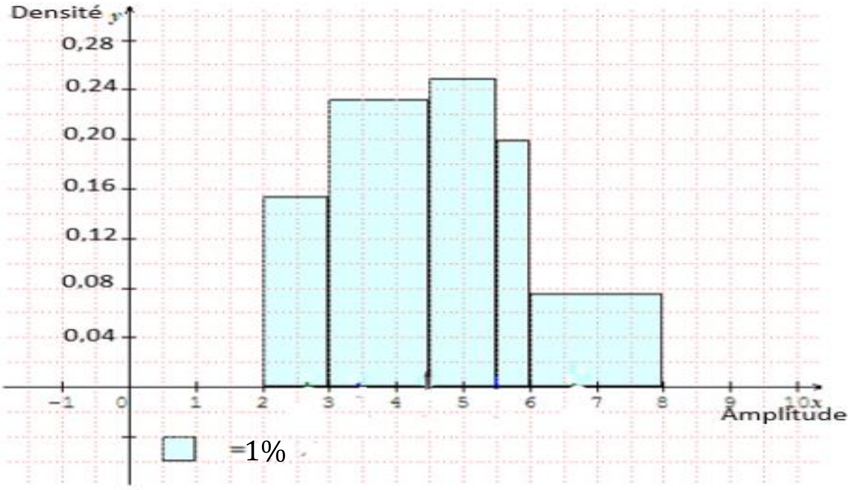
a- Complète le tableau
b- Construis l'histogramme des fréquences de cette série statistique.

Solution

a- Complétons le tableau

Modalité	[2 ; 3[	[3 ; 4,5[	[4,5 ; 5,5[	[5,5 ; 6[	[6 ; 8[
Fréquence	0,15	0,35	0,25	0,1	0,15
Amplitude	1	1,5	1	0,5	2
Centre	2,5	3,75	5	5,75	7
Densité	0,15	0,23	0,25	0,2	0,075

b- Construis l'histogramme des fréquences de cette série statistique



La surface d'un petit carreau est $0{,}5 \times 0{,}02 = 0{,}01$ soit 1%

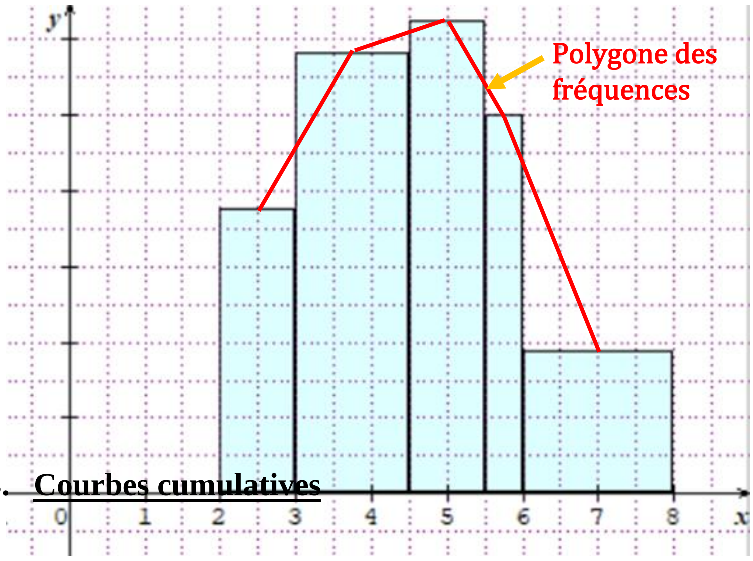
2. Polygones des effectifs et des fréquences

Le polygone des effectifs (resp. fréquences) est obtenu en joignant les milieux successifs des côtés les plus hauts de chaque rectangle de l'histogramme.

Exercice de fixation

Construis le polygone des fréquences de la série statistique dont l'histogramme des fréquences est donné ci-dessous.

Solution



3. Courbes cumulatives

La courbe cumulative des effectifs est la représentation graphique de la fonction F définie sur IR et à valeurs dans l'intervalle [0 ; N] telle que :

- Si $x < x_1$, alors $F(x) = 0$
- Sur chaque intervalle $[x_i; x_{i+1}[$, F coïncide avec la fonction affine g telle que $g(x_i) = N_i$ et $g(x_{i+1}) = N_{i+1}$, où N_i est l'effectif cumulé croissant de $[x_i; x_{i+1}[$ et N_{i+1} celui de $[x_{i+1}; x_{i+2}[$
- Si $x \geq x_{p+1}$, alors $F(x) = N$.

La courbe cumulative des fréquences est la représentation graphique de la fonction F définie sur IR et à valeurs dans l'intervalle [0 ;1] telle que :

- Si$x \leq x_1$, alors $F(x) = 0$
- Sur chaque intervalle $[x_1; x_{i+1}[$ F coïncide avec la fonction affine g telle que $g(x_i)$=F_i et $g(x_{i+1}) = F_{i+1}$, où F_i est la fréquence cumulée croissante de $[x_i; x_{i+1}[$ et F_{i+1} celle de $[x_i ;x_{i+2}[$
- Si $x \geq x_{p+1}$, alors $F(x) = 1$.

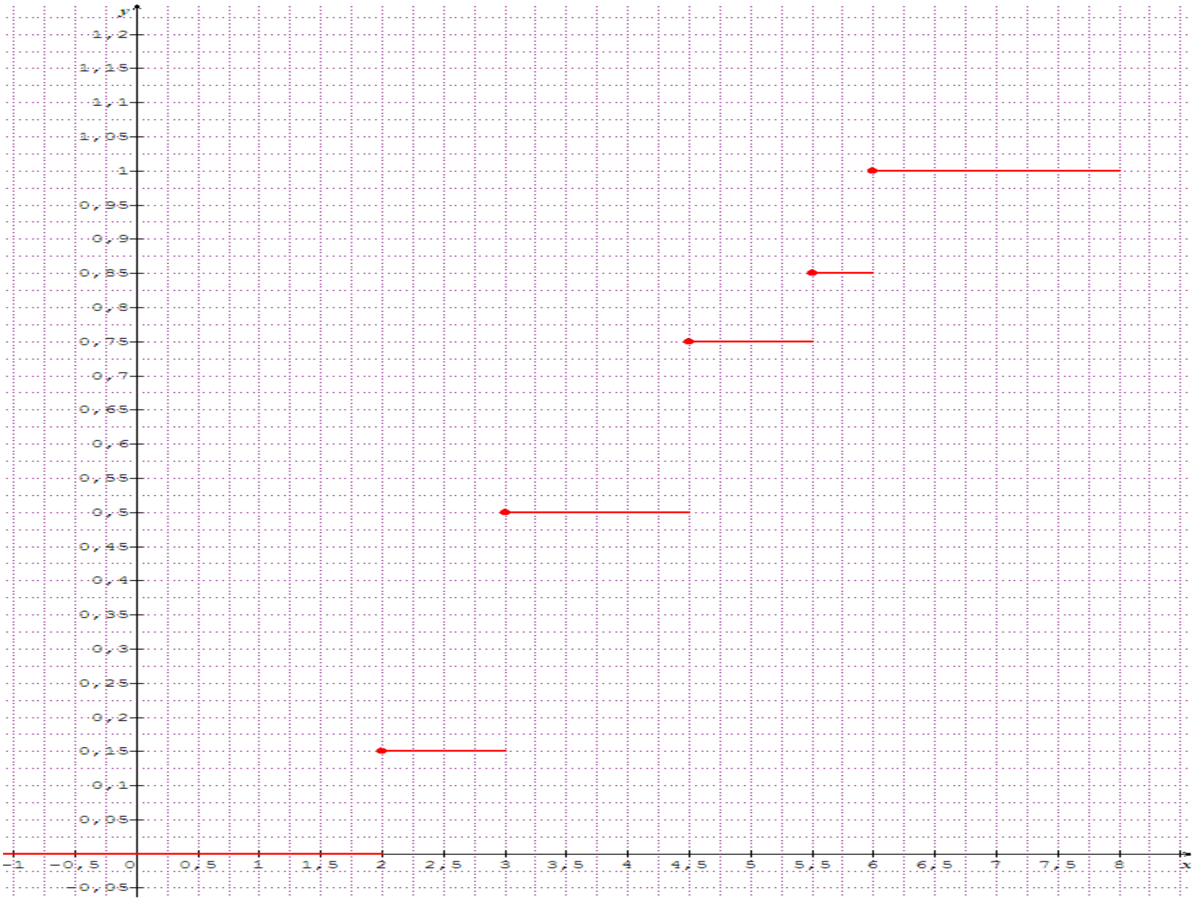
Exercice de fixation

On donne la série statistique suivante :

Classes	[2 ; 3[	[3 ; 4,5[	[4,5 ; 5,5[	[5,5 ; 6[	[6 ; 8[
Fréquences	0,15	0,35	0,25	0,1	0,15
Fréquences cumulées	0,15	0,5	0,75	0,85	1

Construis la courbe cumulative des fréquences de cette série statistique.

Solution



IV. CARACTERISTIQUES DE DISPERSION D'UNE SERIE STATISTIQUE REGROUPEES EN CLASSES

1. Variance

La variance d'une série statistique regroupée en classe, notée V, est donnée par la formule :

$$V = \frac{\sum_{i=1}^{p} n_i\,(c_i - \bar{x})^2}{N} = \frac{1}{N}\left[n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + \cdots + n_p\left(c_p - \bar{x}\right)^2\right]$$

Autre formule :

$$V = \frac{1}{N}\left(\sum_{i=1}^{p} n_i\, {c_i}^2\right) - (\bar{x})^2 = \frac{1}{N}\left(n_1 c_1^2 + n_2 c_2^2 + \cdots + n_p c_p^2\right) - \bar{x}^2$$

Exercice de fixation

On a relevé, pour 125 élèves d'un lycée, le temps consacré à la pratique de sport par semaine. On obtient le tableau suivant :

Temps (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2

a- Justifions que la moyenne est 39,84

b- Calcule la variance de cette série.

Solution

Temps (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[	Total
Effectif	35	41	30	12	5	2	125
Centre	10	30	50	80	120	170	
$n_i\, c_i$	350	1230	1500	960	600	340	4980
$n_i\, {c_i}^2$	3500	36900	75000	76800	72000	57800	322000

a/ la moyenne $\bar{x}$

$$\bar{x}= \frac{4980}{125} = 39{,}84$$

b/ la variance V

$$V=\frac{1}{N}\left(\sum_{i=1}^{6} n_i\, {c_i}^2\right) - (\bar{x})^2 = \frac{322000}{125} - (39{,}84)^2 = 988{,}77$$

2. Ecart type

L'écart type d'une série statistique, noté $\sigma = \sqrt{V}$.

Interprétation

Plus l'écart type est plus élevé, plus la dispersion des valeurs autour de la moyenne est plus élevée

Exercice de fixation

On a relevé, pour 125 élèves d'un lycée, le temps consacré à la pratique de sport par semaine. On obtient le tableau suivant :

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2

Calcule l'écart type de cette série et interprète le résultat.

Solution

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2
Centre	10	30	50	80	120	170

On trouve $V = 988{,}77$

L'écart type σ

$$\sigma = \sqrt{V} = \sqrt{988{,}77} = 31{,}44.$$

Interprétation(

31,44 est élevé donc le temps que les élèves consacrent au sport par semaine dans la majorité n'est pas proche de la moyenne 39,84

3. Ecart absolu moyen

L'écart absolu moyen, noté e_m , est le réel :

$$e_m = \frac{1}{N}\sum_{i=1}^{p} n_i\,|c_i - \bar{x}| = \frac{1}{N}(n_1 \times |c_1 - \bar{x}| + n_2 \times |c_2 - \bar{x}| +\ldots+ n_p \times |c_p - \bar{x}|).$$

Interprétation

L'écart moyen nous indique la distance moyenne entre la moyenne et les valeurs de la série statistique

<u>Exercice de fixation</u>

On a relevé, pour 125 élèves d'un lycée, le temps consacré à la pratique de sport par semaine. On obtient le tableau suivant :

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2

Calcule l'écart absolu moyen de cette série et interprète le résultat.

Solution

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[	**Total**
Effectif	35	41	30	12	5	2	**125**
Centre	10	30	50	80	120	170	
$\boldsymbol{\lvert c_i - \bar{x} \rvert}$	29,84	9,84	10,16	40,16	80,16	130,16	
$\boldsymbol{n_i \lvert c_i - \bar{x} \rvert}$	1044,4	403,44	304,8	481,92	400,8	260,32	**2895,68**

On trouve la moyenne $\bar{x} = 39{,}84$

L'écart absolu moyen

$$e_m = \frac{1}{125}(35 \times |10 - 39{,}84| + 41 \times |30 - 39{,}84| + 30 \times |50 - 39{,}84| + 12 \times |80 - 39{,}84| + 5 \times |120 - 39{,}84| + 2 \times |170 - 39{,}84|)$$

$$e_m = \frac{2895{,}68}{125} = 23{,}17$$

<u>Interprétation.</u>

Le temps consacré à la pratique de sport par semaine d'un élève est à 23,17 min de la moyenne.

4. <u>Ecart interquartile</u>

L'écart interquartile est la différence entre le troisième et le premier quartile. C'est le nombre $Q_3 - Q_1$.

<u>Interprétation</u>

50% des valeurs sont comprises entre le premier quartile et le troisième quartile .

<u>Exercice de fixation</u>

On a relevé, pour 125 élèves d'un lycée, le temps consacré à la pratique de sport par semaine. On obtient le tableau suivant :

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2

Calcule l'écart interquartile de cette série et interprète le résultat.

Solution

Temps en (min)	[0 ; 20[	[20 ; 40[	[40 ; 60[	[60 ; 100[	[100 ; 140[	[140 ; 200[
Effectif	35	41	30	12	5	2
Effectifs cumulés croissants	35	76	106	118	123	125

- On a $125 \times \frac{25}{100} = 31{,}25$.

31,25 est plus petit que la 35^{e} valeur. Donc $Q_1 \in [0\,;\,20[$.

$\frac{Q_1-0}{20-0} = \frac{31{,}25-0}{35-0}$ donc $\boldsymbol{Q_1 = 17{,}86}$

- On a $125 \times \frac{75}{100} = 93{,}75$.

93,75 est compris entre la 76^{e} et la 106^{e}valeur. Donc $Q_3 \in [40\,;\,60[$.

$\frac{Q_3-40}{60-40} = \frac{93{,}75-76}{106-76}$ donc $Q_3 = 51{,}83$

L'écart interquartile est $Q_3 - Q_1$ =51,83-17,86=33,97

Interprétation

L'étendue est 200.

L'interquartile est environ égale à 25% de l'étendue donc 50% des élèves ont leur temps consacré à la pratique du sport proche de la médiane.

C. SITUATION COMPLEXE

Le président du comité de gestion scolaire (COGES) de ton lycée veut acquérir une machine pour stocker sa récolte dans des sacs de 50 kg. Un commerçant lui propose deux machines, A et B, qu'il teste sur 80 sacs. Le tableau ci-dessous indique les résultats obtenus.

Classe	Machine A	Machine B	Classe	Machine A	Machine B
$[49; 49{,}2[$	7	4	$[49{,}8; 50[$	11	14
$[49{,}2; 49{,}4[$	6	6	$[50; 50{,}2[$	14	16
$[49{,}4; 49{,}6[$	14	9	$[50{,}2; 50{,}4[$	9	17
$[49{,}6; 49{,}8[$	14	11	$[50{,}4; 50{,}6[$	5	3

Un agent de l'agriculture du service qualité estime que la machine est bonne si les conditions suivantes sont simultanément réalisées :

- la moyenne $\bar{x}$ doit être comprise entre 49,7 et 50,3 ;
- l'écart-type σ doit être inférieur à 0,5 kg ;
- l'intervalle $[49{,}3; 50{,}5]$ doit contenir 85% des sacs.

Sollicité par un membre du bureau du COGES pour le choix de la machine convenable, tu décides de répondre à la préoccupation de leur président en utilisant tes connaissances en mathématiques.

Solution

Pour vérifier la qualité de cette machine je vais utiliser la leçon sur les statistiques .

Pour cela je vais :

- Dresser le tableau des effectifs, des fréquences et des fréquence cumulées croissantes de chaque machine.
- Calculer le stockage moyen de chaque machine.
- Calculer la variance l'écart type de chaque série statistique.
- Utiliser l'interpolation linéaire pour donner une estimation de la fréquence correspondante à la valeur 49,3 et 50,5.
    - Le tableau des effectifs, des fréquences et des fréquence cumulées croissantes de chaque machine.

Machine A

Stockage	[49; 49,2[	[49,2;49,4[	[49,4;49,6[	[49,6; 49,8[	[49,8; 50[	[50; 50,2[	[50,2; 50,4[	[50,4; 50,6[
Nombre de sacs	7	6	14	14	11	14	9	5
Centres	49,1	49,3	49,5	49,7	49,9	50,1	50,3	50,5
Fréquence (%)	8,8	7,5	17,5	17,5	13,8	17,5	11,3	6,3
Fréquence cumulées croissantes(%)	8,8	16,3	33,8	51,3	65	82,5	93,8	100

Machine B

Stockages	[49; 49,2[	[49,2;49,4[	[49,4;49,6[	[49,6; 49,8[	[49,8; 50[	[50; 50,2[	[50,2; 50,4[	[50,4; 50,6[
Nombre de sacs	4	6	9	11	14	16	17	3
Centres	49,1	49,3	49,5	49,7	49,9	50,1	50,3	50,5
Fréquence (%)	5	7,5	11,3	13,8	17,5	20	21,3	3,8
Fréquence cumulées croissantes(%)	5	12,5	23,8	37,5	55	75	96,3	100

- Le stockage moyen de chaque machine.

Machine A :

$$\boldsymbol{\bar{x}} = \frac{\sum_{i=1}^{8} n_i c_i}{N} = \frac{3983,8}{80} = 49,79$$

Machine B :

$$\boldsymbol{\bar{x}} = \frac{\sum_{i=1}^{8} n_i c_i}{N} = \frac{3991,2}{80} = 49,89$$

- La variance et l'écart type de chaque série statistique.

Machine A

$$V=\frac{1}{N}\left(\sum_{i=1}^{8} n_i\, {c_i}^2\right) - (\bar{x})^2 = \frac{198395,7}{80} - (49,78)^2 = 0,39$$

$$\sigma = \sqrt{0,39} = 0,62$$

Machine B

$$V=\frac{1}{N}\left(\sum_{i=1}^{8} n_i\, {c_i}^2\right) - (\bar{x})^2 = \frac{199132}{80} - (49,89)^2 = 0,14$$

$$\sigma = \sqrt{0,14} = 0,37$$

- Une estimation de la fréquence correspondante à la valeur 49,3 et 50,5par l'interpolation linéaire.

Machine A

A l'aide du tableau des fréquences cumulées croissantes , On a :

49,2	49,3	49,4
8,8	b	16,3

Machine B

A l'aide du tableau des fréquences cumulées croissantes , On a :

49,2	49,3	49,4
5	b	12,5

$$\frac{49,3-49,2}{49,4-49,2}=\frac{b-8,8}{16,3-8,8}$$

$\frac{b-8,8}{7,5}=\frac{0,1}{0,2}$ donc b =12,55

50,4	50,5	50,6
93,8	b	100

$$\frac{50,5-50,4}{50,6-50,4}=\frac{b-93,8}{100-93,8}$$

$\frac{b-93,8}{6,2}=\frac{0,1}{0,2}$ donc b =96,9

La fréquence de [49,3;50,5] est
$96,9-12,55=84,35\%$

$$\frac{49,3-49,2}{49,4-49,2}=\frac{b-5}{12,5-5}$$

$\frac{b-5}{7,5}=\frac{0,1}{0,2}$ donc b =8,75

50,4	50,5	50,6
96,3	b	100

$$\frac{50,5-50,4}{50,6-50,4}=\frac{b-96,3}{100-96,3}$$

$\frac{b-96,3}{3,7}=\frac{0,1}{0,2}$ donc b =98,15

La fréquence de [49,3;50,5] est
$98,15-8,75=89,4\ \%$

Après calcul, on remarque que :

Pour la machine A on a :

- la moyenne calculée (49,8) est comprise entre 49,7 et 50,3 ;
- l'écart-type calculé (0,39) est inférieur à 0,5
- l'intervalle [49,3;50,5] contient 84,35% des sacs donc la machine n'arrive pas à contenir 85% des sacs.

Pour la machine B on a :

- la moyenne calculée (49,89) est comprise entre 49,7 et 50,3 ;
- l'écart-type calculé (0,37) est inférieur à 0,5

l'intervalle [49,3;50,5] contient 89,4% des sacs donc la machine arrive à contenir 85% des sacs.

Donc le président du comité de gestion scolaire (COGES) devra choisir la machine B.

D. EXERCICES

Exercice 1

Le tableau ci- dessous indique la répartition du salaire journalier (en milliers de francs CFA) des ouvriers d'une entreprise.

Salaire	[1; 2[	[2; 3[	[3; 5[	[5; 8[
Nombre d'ouvriers	34	76	51	39

1. Dans un tableau détermine la densité et l'amplitude de cette série statistique.
2. Construis l'histogramme des effectifs de cette série statistique.
3. Construis le polygone des effectifs cumulés croissants et celui des polygones des effectifs cumulés décroissants.
4. a) Déduis graphiquement une valeur approchée de la médiane.
   b) Détermine par le calcul une estimation de la médiane.

Solution

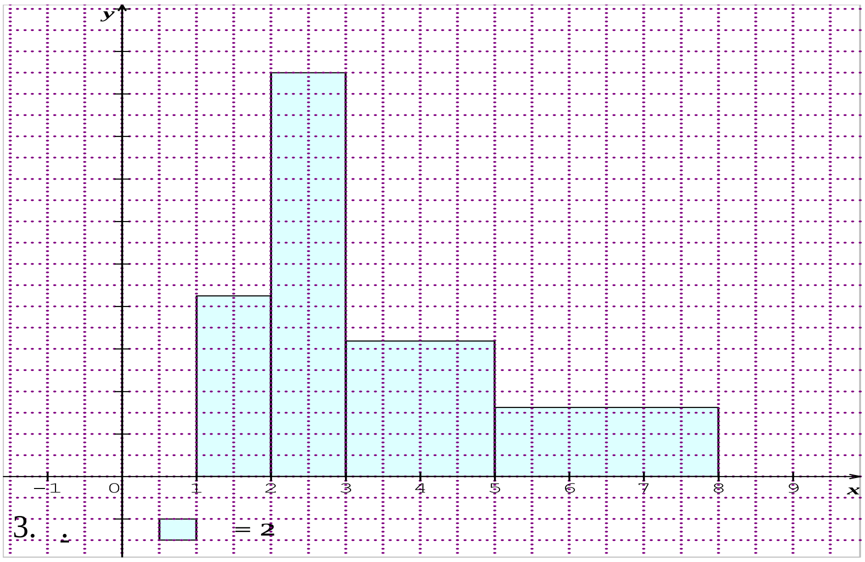
1. Tableau des fréquences

Salaire	[1; 2[	[2; 3[	[3; 5[	[5; 8[	Total
Nombre d'ouvriers	34	76	51	39	200
Amplitude	2	1	2	3	
Densité	17	76	25,5	13	

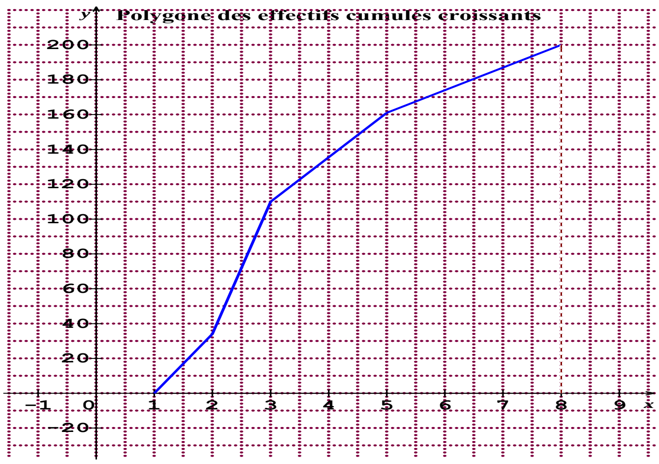
2. L'histogramme des effectifs

Echelle ; un petit carreau correspond à 2 ouvriers soit $1cm^2 \rightarrow 8$ ouvriers

Salaire	[1; 2[	[2; 3[	[3; 5[	[5; 8[	Total
Nombre d'ouvriers	34	76	51	39	200
Largeur	1	1	2	3	
hauteur	4,25	9,5	3,2	1,6	



3.

Le polygone des effectifs cumulés croissants	Le polygone des effectifs cumulés décroissants.
	

Exercice 2

On donne la série statistique suivante :

Tranche d'âge (en année)	$[0;10[$	$[10;15[$	$[15;17[$	$[17;20[$
Effectif (n_i)	7	16	18	9

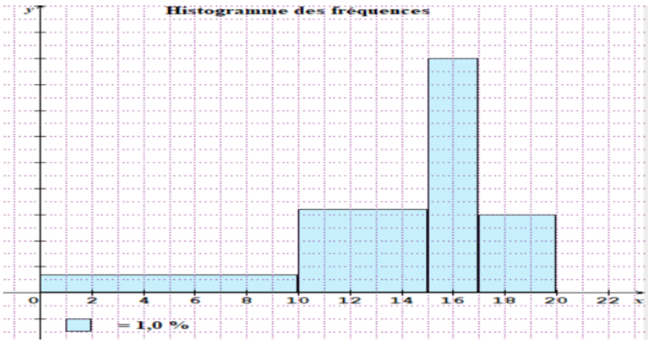
1. Dresse, en pourcentage, le tableau des fréquences cumulées croissantes de cette série.
2. Construis l'histogramme des fréquences de cette série statistique.
3. Construis le polygone des fréquences cumulées croissantes.
4. a) Déduis graphiquement une valeur approchée à 10^{-1} près des quartiles de cette série statistique.
   b) Détermine par le calcul une estimation des quartiles de cette série statistique.
   (On donnera les résultats à l'arrondi d'ordre 1).
   c) Calcul l'écart interquartile.

<u>Solution</u>

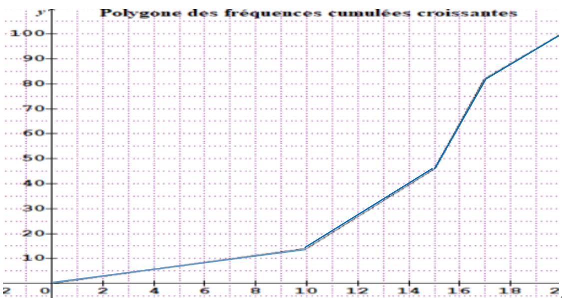
1. Le tableau des fréquences cumulées croissantes de cette série en pourcentage.

Tranche d'âge (en année)	$[0;10[$	$[10;15[$	$[15;17[$	$[17;20[$
Effectif (n_i)	7	16	18	9
Fréquence (%)	14	32	36	18
Fréquence cumulées croissantes(%)	14	46	82	1

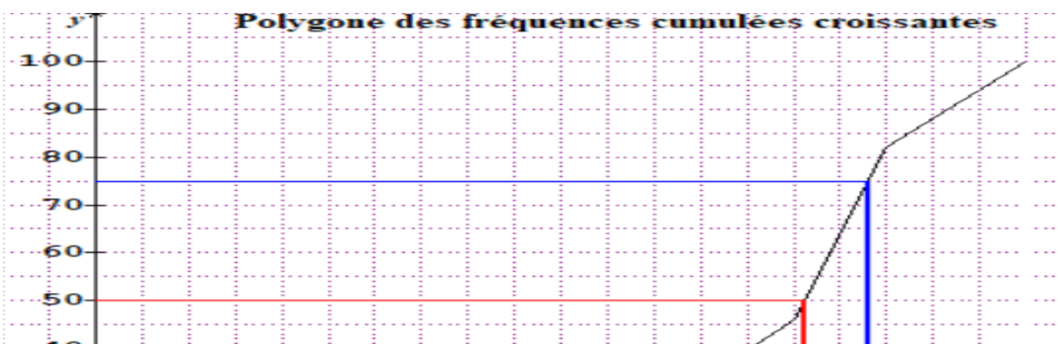
2. Construis l'histogramme des fréquences de cette série statistique.



3. Construis le polygone des fréquences cumulées croissantes.



4. a) Une valeur approchée à 10^{-1} près des quartiles de cette série statistique.



Q_1 Med Q_3

$\boldsymbol{Q_1 = 11,7}$ $\boldsymbol{Q_3 = 16,6}$

b) Une estimation des quartiles par le calcul de cette série statistique.

Tranche d'âge (en année)	$[0; 10[$	$[10; 15[$	$[15; 17[$	$[17; 20[$
Fréquence (%)	14	32	36	18
Fréquence cumulées croissantes(%)	14	46	82	1

• On a $100 \times \frac{25}{100} = 25$ 25 est compris entre la 14^e et la 46^evaleur. Donc $Q_1 \in [10\,;\,15[$. $\frac{Q_1-10}{15-10} = \frac{25-14}{46-14}$ donc $\boldsymbol{Q_1 = 11,72}$	• On a $100 \times \frac{75}{100} = 75$. 75 est compris entre la 46^evaleur et la 82^evaleur. Donc $Q_3 \in [15\,;\,17[$ $\frac{Q_3-15}{17-15} = \frac{75-46}{82-46}$ donc $Q_3 = 16{,}61$

c) Calcul l'écart interquartile.

L'écart interquartile est $Q_3 - Q_1$ =16,61-11,72=4,89

<u>Exercice 3</u>

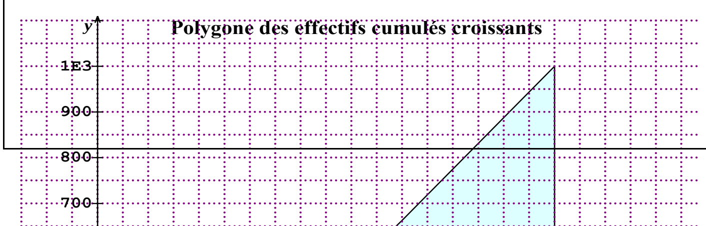
On considère la série statistique présentée dans le tableau suivant :

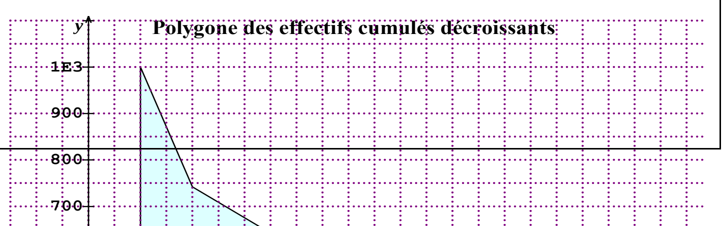
Classe	$[1; 2[$	$[2; 4[$	$[4; 5[$	$[5; 6[$	$[6; 9[$
Effectif	258	133	164	108	337

1. Construis les polygones des effectifs cumulés croissants et décroissants.
2. Détermine le nombre d'individus dont la modalité est :
   a) inférieure à 3 ; b) supérieure à 5,3 ; c) comprise entre 3 et 5,3
3. Calcule la moyenne, la variance et l'écart type de cette série.

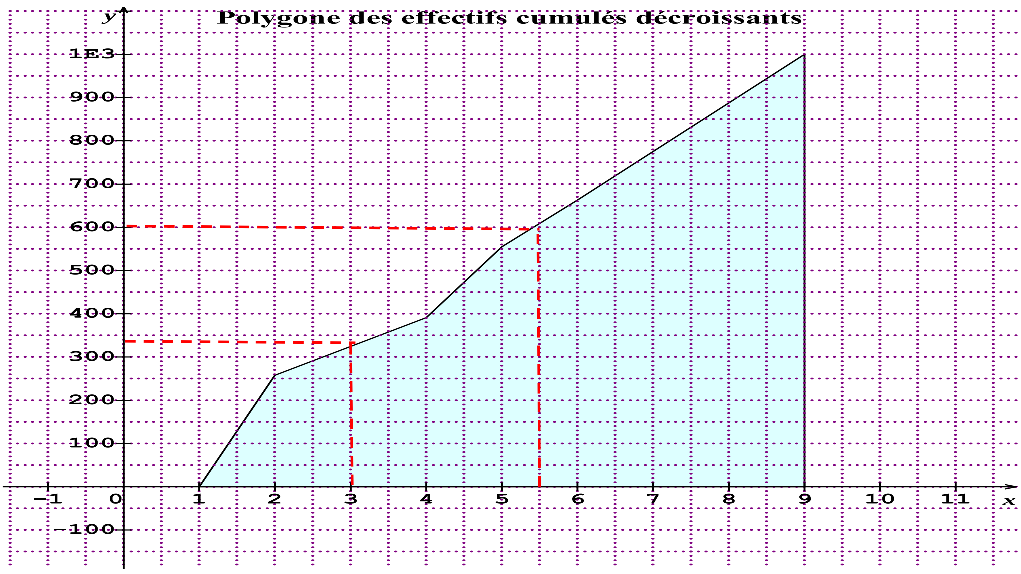
<u>Réponse</u>

1. :





1. Le nombre d'individus dont la modalité est :



- inférieure à 3

Graphiquement on a 340 individus qui ont une modalité inférieure 3

- supérieure à 5,3

Graphiquement on a 600 individus qui ont une modalité inférieure 3

- comprise entre 3 et 5,3

On a 600-340=260 individus.

2.

Classe	[1; 2[	[2; 4[	[4; 5[	[5; 6[	[6; 9[	Total
Effectif	258	133	164	108	337	1000
Centre	1,5	3	4,5	5,5	7,5	
$\boldsymbol{n_i\, c_i}$	387	399	738	594	2527,5	4645,5
$\boldsymbol{n_i\, c_i^{\,2}}$	580,5	1197	3321	3267	18956,25	27321,75

la moyenne,

$$\bar{\boldsymbol{x}} = \frac{\sum_{i=1}^{5} n_i c_i}{N} = \frac{4645,5}{1000} = 4,65$$

La variance

$$V=\frac{1}{N}\sum_{i=1}^{5} n_i\, c_i{}^2) - (\bar{x})^2 = \frac{27321,75}{1000} - (4,65)^2 = 5,7$$

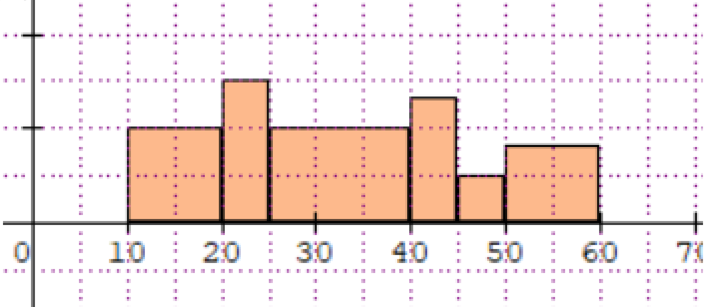
L’écart type.

$$\sigma = \sqrt{5,7} = 2,4$$

Exercice 4

L’histogramme ci-contre représente les recettes en milliers de francs CFA de 300 commerçants. $1cm^2$ représente 60 individus.

1. Justifie qu’on obtient le tableau des effectifs suivant

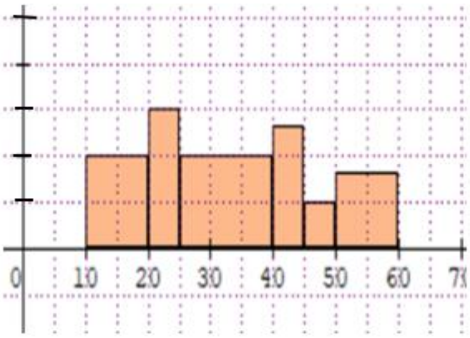


Classes	$[10;20[$	$[20;25[$	$[25;40[$	$[40;45[$	$[45;50[$	$[50;60[$
Effectifs	60	45	90	40	15	50

2. Détermine l’écart –type de cette série statistique. Interprète ce résultat.
3. Détermine l’écart interquartile de cette série statistique. Interprète ce résultat.

Solution

1.



On sait que l’aire d’un rectangle est hauteur × base

Classes	$[10;20[$	$[20;25[$	$[25;40[$	$[40;45[$	$[45;50[$	$[50;60[$
Aire du rectangle(cm^2)	2	1,5	3	1,33	0,5	1,67

Pour trouver l’aire de la classe [50 ;60[on a la base est égale à 1cm et la hauteur est égale environ à 1,67.

On obtient le tableau des effectifs

Classes	$[10;20[$	$[20;25[$	$[25;40[$	$[40;45[$	$[45;50[$	$[50;60[$

Effectifs	60	45	90	40	15	50

2. L'écart –type de cette série statistique.

Classe	$[10;20[$	$[20;25[$	$[25;40[$	$[40;45[$	$[45;50[$	$[50;60[$	Total
Effectif	60	45	90	40	15	60	300
Centre	15	22,5	32,5	42,5	47,5	55	
$\boldsymbol{n_i\, c_i}$	900	1012,5	2925	1700	712,5	2750	10000
$\boldsymbol{n_i\, c_i^{\,2}}$	13500	22781,25	95062,5	72250	33843,75	151250	388687,5

la moyenne : $\boldsymbol{\bar{x}} = \frac{\sum_{i=1}^{6} n_i c_i}{N} = \frac{10000}{300} = 33{,}33$

La variance : V=$\frac{1}{N}$ $(\sum_{i=1}^{6} n_i\, c_i^{\,2}) - (\bar{x})^2 = \frac{388687{,}5}{300}$ - $(33{,}33)^2$ = 184,74

L'écart type : $\sigma = \sqrt{184{,}74} = 13{,}59$

<u>Interprétation</u>

13,59 est élevé donc les recettes sont dispersées autour de la moyenne 33,33.

3. L'écart interquartile de cette série statistique.

Classes	$[10;20[$	$[20;25[$	$[25;40[$	$[40;45[$	$[45;50[$	$[50;60[$
Effectifs	60	45	90	40	15	50
Effectifs cumulées croissantes	60	105	195	235	250	300

- On a 300$\times \frac{25}{100} = 75$

75 est compris entre la 60^e et la 105^evaleur. Donc $Q_1 \in [20\,;\,25[$.

$\frac{Q_1-20}{25-20} = \frac{75-60}{105-60}$ donc $\boldsymbol{Q_1 = 21,67}$

- On a 300$\times \frac{75}{100} = 225$

225 est compris entre la 195^e et la 235^evaleur. Donc $Q_3 \in [40\,;\,45[$.

$\frac{Q_1-40}{45-40} = \frac{225-195}{235-195}$ donc $\boldsymbol{Q_3 = 43,75}$

$Q_3 - Q_1 = 22,08$

Interprétation

50% des recettes sont dispersées dans un intervalle d'amplitude 22,08 qui contient la médiane

Exercice 5

On considère la série statistique suivante :

Surface (en m²)	[0; 10[	[10; 15[	[15; 17[	[17 ; 20[
Individu (x_i)	20	10	15	5

1. Détermine la classe modale et calcule la moyenne $\bar{x}$ de cette série.
2. Calcule les caractéristiques de dispersion de cette série.

Solution

1. Déterminons la densité de chaque intervalle.

Surface (en m²)	[0; 10[	[10; 15[	[15; 17[	[17 ; 20[
Individu (x_i)	20	10	15	5
Amplitude	10	5	2	3
Densité	2	2	7,5	1,67

La classe modale est : [15; 17[

2. Les caractéristiques de dispersions.

Surface (en m²)	[0; 10[	[10; 15[	[15; 17[	[17 ; 20[	Total
Individu (x_i)	20	10	15	5	50
Centre	5	12,5	16	18,5	
$\boldsymbol{n_i\, c_i}$	100	125	240	92,5	557,5
$\boldsymbol{n_i\, {c_i}^2}$	500	1562,5	3840	1711,25	7613,75
$\boldsymbol{\lvert c_i - \bar{x}\rvert}$	6,15	1,35	4,85	7,35	
$\boldsymbol{n_i \lvert c_i - \bar{x}\rvert}$	123	13,5	72,75	36,75	246

la moyenne : $\boldsymbol{\bar{x}} = \frac{\sum_{i=1}^{4} n_i c_i}{N} = \frac{557,5}{50} = 11,15$

La variance : V$=\frac{1}{N}\left(\sum_{i=1}^{4} n_i\, {c_i}^2\right) - (\bar{x})^2 = \frac{7613,75}{50}$ - $(11,15)^2$ = 27,95

L'écart type : $\sigma = \sqrt{27,95} = 5,29$

L'écart absolu moyen.

$e_m = \frac{1}{N}\sum_{i=1}^{4} n_i \lvert c_i - \bar{x}\rvert = \frac{246}{50} = 4,92$

L’écart ’interquartile

Surface (en m²)	$[0; 10[$	$[10; 15[$	$[15; 17[$	$[17 ; 20[$
Individu (x_i)	20	10	15	5
Effectifs cumulées croissantes	20	30	45	50

- On a 50$\times \frac{25}{100} = 12{,}5$

12,5 est plus petite que la 20^evaleur. Donc $Q_1 \in [0 ; 10[$.

$\frac{Q_1}{10} = \frac{12{,}5}{20}$ donc $Q_1 = 6{,}25$

- On a 50$\times \frac{75}{100} = 37{,}5$

37,5 est compris entre la 30^e et la 45^evaleur. Donc $Q_3 \in [15 ; 17[$.

$\frac{Q_1 - 15}{17-15} = \frac{37{,}5-30}{45-30}$ donc $Q_3 = 16$

L’interquartile est donc :

$Q_3 - Q_1 = 9{,}75$