

Brevet Professionnel

Mathématiques

Module Statistique et probabilités

Statistique à une variable

Partie 2

Comparer des séries statistiques
à l'aide d'indicateurs

Nom :

Groupe :

Caractéristiques de position

Activités

Voici les séries de notes obtenues par 3 élèves :

Jérôme : 4 ; 6 ; 18 ; 7 ; 17 ; 12 ; 12 ; 18

Bertrand : 13 ; 13 ; 12 ; 10 ; 12 ; 3 ; 14 ; 12 ; 14 ; 15

Julie : 15 ; 9 ; 14 ; 13 ; 10 ; 12 ; 12 ; 11 ; 10

1. Moyennes

$$M(\text{Jérôme}) = (4 + 6 + 18 + 7 + 17 + 12 + 12 + 18) : 8 \approx 11,8$$

$$M(\text{Bertrand}) = (13 + 13 + 12 + 10 + 12 + 3 + 14 + 12 + 14 + 15) : 10 = 11,8$$

$$M(\text{Julie}) = (15 + 9 + 14 + 13 + 10 + 12 + 12 + 11 + 10) : 9 \approx 11,8$$

La moyenne est une caractéristique de position.

2. Médianes

Définition :

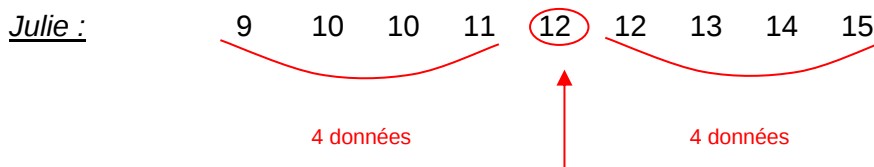
La **médiane** m est **une valeur de la série** telle que la **moitié** de l'effectif ait des valeurs inférieures à m , l'autre **moitié** des valeurs supérieures à m .

Pour déterminer les notes médianes, il faut **ordonner** les séries. La médiane partage l'effectif en deux.





$$m(\text{Bertrand}) = (12 + 13) : 2 = 12,5$$



$$m(\text{Julie}) = 12$$

La médiane est une caractéristique de position.

3. Moyenne dans le cas de regroupement par classe (caractère continu)

Ex : Tailles des élèves de 2^{nde}5 en cm :

174 – 160 – 161 – 166 – 177 – 172 – 157 – 175 – 162 – 169 – 160 – 165 – 170 – 152 – 168 156 –
163 – 167 – 169 – 158 – 164 – 151 – 162 – 166 – 156 – 165 – 179

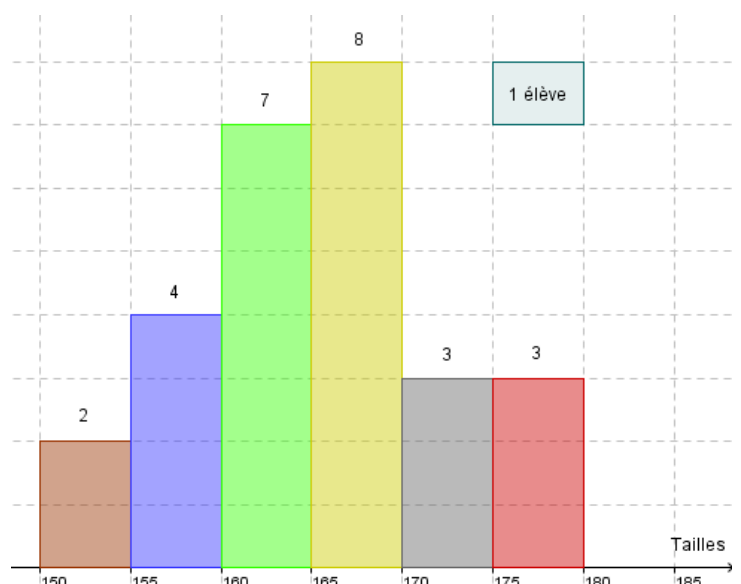
Regrouper cette série de tailles par classes de longueur 5 cm et calculer les fréquences en %
arrondies à l'unité :

Tailles	$150 \leq t < 155$	$155 \leq t < 160$	$160 \leq t < 165$	$165 \leq t < 170$	$170 \leq t < 175$	$175 \leq t < 180$
Effectifs						
Fréquences						

L'effectif total est

HISTOGRAMME

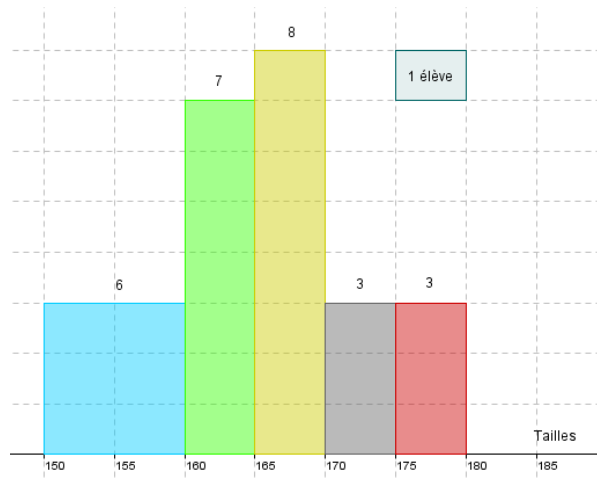
DES EFFECTIFS DES TAILLES



Remarque :

Dans un histogramme, l'aire des rectangles est proportionnelle à l'effectif (ou à la fréquence).

Ainsi, dans l'exemple, en regroupant les deux premières classes, on obtiendrait la représentation suivante :



Calcul de la moyenne :

a) Calcul de la moyenne en centrant les classes :

- On calcule le centre de chaque classe :

Classes centrées	152	157				
Effectifs	2	4				

- Il s'agit la moyenne pondérée.

Définition :

La **moyenne** d'une série statistique dont les valeurs sont $x_1, x_2, \dots, x_k$ et les effectifs correspondants $n_1, n_2, \dots, n_k$ est notée $\bar{x}$ et est égale à $\bar{x} = \frac{n_1 x_1 + \dots + n_k x_k}{n_1 + \dots + n_k}$.

Ainsi dans l'exemple :

$$(152 \times 2 + 157 \times 4 + 162 \times 7 + 167 \times 8 + 172 \times 3 + 177 \times 3) : 27$$

$$= 4449 : 27 \approx 164,8 \text{ cm}$$

b) Calcul de la moyenne exacte :

$$(174 + 160 + 161 + 166 + 177 + 172 + 157 + 175 + 162 + 169 + 160 + 165 + 170 + 152 + 168 + 156 + 163 + 167 + 169 + 158 + 164 + 151 + 162 + 166 + 156 + 165 + 179) : 27$$

$$= 4444 : 27 \approx 164,6 \text{ cm}$$

La méthode de calcul de moyenne en centrant les classes est très fiable (*ici : 2 mm d'erreur*)

4. Mode

Exemple 1 : **Nombre de personnes par voiture.**



Relever la valeur du caractère étudié la plus représentée :

⇒ La valeur du caractère la plus représentée (pour laquelle la est maximale) est le

Exemple 2 : Tailles des élèves d'une classe en cm (voir paragraphe 3)

Repérer la classe pour laquelle l'effectif est le plus important :

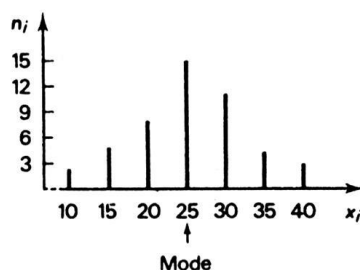
⇒ Il s'agit de la

Calculer la moyenne des valeurs des bornes de la classe :

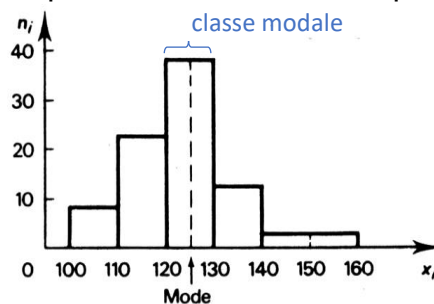
⇒ Il s'agit du

Définition

Le **mode** d'une série statistique à *caractère qualitatif ou quantitatif discret* est la **valeur du caractère pour laquelle l'effectif est le plus élevé.**



La **classe modale** d'une série statistique à *caractère quantitatif continu* est la classe qui correspond à l'effectif le plus important. Le **mode** correspond alors au **centre de classe.**



Caractéristiques de dispersion

1. Etendue

Définition :

L'**étendue** d'une série statistique est la **différence** entre la plus grande valeur et la plus petite valeur de la série.

En reprenant les données de l'exemple étudié au paragraphe I. :

$$E_{(Jérôme)} = 18 - 4 = 14$$

$$E_{(Bertrand)} = 15 - 3 = 12$$

$$E_{(Julie)} = 15 - 9 = 6$$

Remarque « 3 » est négligeable dans la série de Bertrand. On peut alors **élaguer** la série et prendre 10 comme minimum pour calculer l'étendue :

$$E_{(Bertrand)} = 15 - 10 = 5 \quad (\text{élaguée})$$

L'étendue est une caractéristique de dispersion.

2. Quartiles

Définitions :

Le **premier quartile Q1** est la **plus petite valeur** de la série telle qu'**au moins 25 %** des autres valeurs de la série sont inférieures ou égales à cette valeur.

Le **troisième quartile Q3** est la **plus petite valeur** de la série telle qu'**au moins 75 %** des autres valeurs de la série sont inférieures ou égales à cette valeur.

- ⇒ Pour déterminer les quartiles, il faut **ordonner** les séries.
- ⇒ Le premier quartile est la donnée de la série se trouvant « au » quart de l'effectif.
- ⇒ Le troisième quartile est la donnée de la série se trouvant « au » trois-quarts de l'effectif.

Les quartiles sont des caractéristiques de dispersion.

Jérôme : 4 6 7 12 12 17 18 18

$\frac{1}{4} \times 8 = 2$, le premier quartile est la 2^e donnée de la série ordonnée.

$\frac{3}{4} \times 8 = 6$, le troisième quartile est la 6^e donnée de la série ordonnée.

$$Q_1(\text{Jérôme}) = 6$$

$$Q_3(\text{Jérôme}) = 17$$

Bertrand : 3 10 12 12 12 13 13 14 14 15

$\frac{1}{4} \times 10 = 2.5$, le premier quartile est la 3^e donnée de la série ordonnée.

$\frac{3}{4} \times 10 = 7.5$, le troisième quartile est la 8^e donnée de la série ordonnée.

$$Q_1(\text{Bertrand}) = 12$$

$$Q_3(\text{Bertrand}) = 14$$

Julie : 9 10 10 11 12 12 13 14 15

$\frac{1}{4} \times 9 = 2.25$, le premier quartile est la 3^e donnée de la série ordonnée.

$\frac{3}{4} \times 9 = 6.75$, le troisième quartile est la 7^e donnée de la série ordonnée.

$$Q_1(\text{Julie}) = 10$$

$$Q_3(\text{Julie}) = 13$$

Interprétations des différentes caractéristiques (exemple des notes des 3 élèves)

$M(\text{Jérôme}) = 11,8$ $m(\text{Jérôme}) = 12$ $E(\text{Jérôme}) = 14$ $Q_1(\text{Jérôme}) = 6$ $Q_3(\text{Jérôme}) = 17$

$M(\text{Bertrand}) = 11,8$ $m(\text{Bertrand}) = 12,5$ $E(\text{Bertrand}) = 5$ $Q_1(\text{Bertrand}) = 12$ $Q_3(\text{Bertrand}) = 14$

$M(\text{Julie}) \approx 11,8$ $m(\text{Julie}) = 12$ $E(\text{Julie}) = 6$ $Q_1(\text{Julie}) = 10$ $Q_3(\text{Julie}) = 13$

Les moyennes sont environ égales et pourtant les notes ne se répartissent pas de la même manière autour de cette caractéristique de position. Les étendues sont très différentes.

Dire que Jérôme a une médiane égale à 12 signifie que Jérôme a obtenu autant de notes au dessus de 12 que de notes en dessous de 12.

Dire que le premier quartile de Bertrand est égal à 12 signifie qu'au moins un quart des notes de Bertrand sont inférieures à 12.

Dire que le troisième quartile de Julie est égal à 13 signifie qu'au moins trois quarts des notes de Julie sont inférieures à 13.

3. Ecart Interquartile

Définitions :

L'**écart interquartile** est une mesure de dispersion qui s'obtient en faisant la différence entre le troisième et le premier quartile :

$$EI = Q_3 - Q_1$$

Exercice

La grille des salaires des employés d'une PME est donnée par le tableau ci-dessous :

Catégorie	Ouvrier simple	Ouvrier qualifié	Cadre moyen	Cadre supérieur	Dirigeant
Effectif	40	19	12	8	1
Salaire en euros	950	1300	1700	3500	8000

1. Quel est l'effectif total de l'entreprise ? Calculer le salaire moyen.

.....

.....

2. Quel est le salaire médian pour cette entreprise ?

.....

.....

3. Quels sont les premier et troisième quartiles ?

.....

.....

.....

4. Couple médiane – quartiles : représentation graphique

Définitions

On appelle **diagramme en boîte** (ou **diagramme à moustaches**) un diagramme résumant les principaux paramètres de la série tel celui-ci :

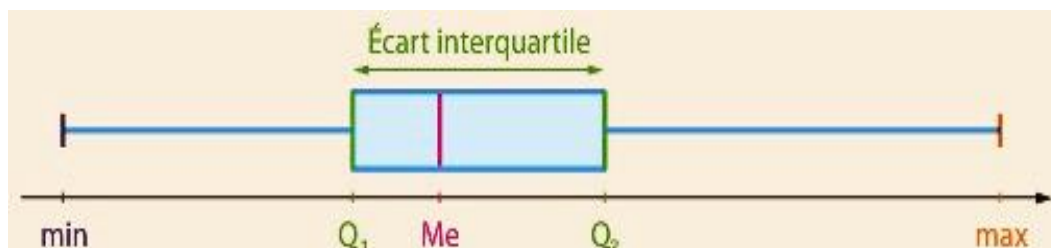


FIGURE 1 – diagramme en boîte, ou encore diagramme de Tuckey

Remarques

Ne pas oublier de faire figurer un axe gradué!

Les diagrammes en boîtes sont très utiles pour comparer rapidement des séries statistiques.

Étude d'un exemple

Lors d'une enquête on a demandé à 202 lycéens le montant mensuel de leur argent de poche. Les résultats ont été regroupés dans le tableau suivant.

Argent de poche (€)	20	50	70	90	110	140
Effectif	12	22	50	64	42	12
ECC	12	34	84	148	190	202

Méthode

Utilisation de la calculatrice On peut utiliser la calculatrice ^a pour déterminer les principaux paramètres d'une série statistique. Le principe est de saisir les valeurs dans une liste, et les effectifs dans une autre liste (si la série est donnée valeur par valeur, on n'utilise qu'une seule liste!). Il est aussi possible d'afficher le diagramme en boîte.

a. voir par exemple sur le site : <http://www.mathslycee.fr/2013-09-15-16-16-29/premiere-s/132-statistiques-et-probabilites/statistiques-descriptives/109-1s-sp1-stats-et-calculatrices>

L1	L2	L3	2	EDIT	TESTS	Var Stats L1,L2	1-Var Stats
20	12	-----		1:1-Var Stats			n=202
50	22			2:2-Var Stats			minX=20
70	50			3:Med-Med			Q1=70
90	64			4:LinReg(ax+b)			Med=90
110	42			5:QuadReg			Q3=110
140	12			6:CubicReg			maxX=140
-----				7:QuartReg			
L2(?) =							

FIGURE 2 – calcul des paramètres avec une calculatrice TI

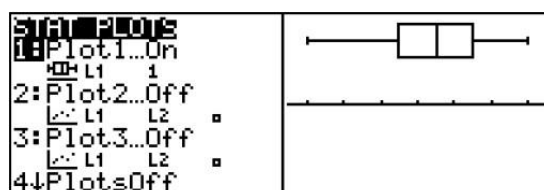


FIGURE 3

Déterminer la médiane :

.....

.....

.....

.....

Déterminer le premier quartile :

.....

.....

.....

.....

Déterminer le troisième quartile :

.....

.....

.....

.....

Paramètres de dispersion : Déterminer l'étendue et l'interquartile

.....

.....

.....

Exercice en ligne – diagramme moustache <https://learningapps.org/view12660205>

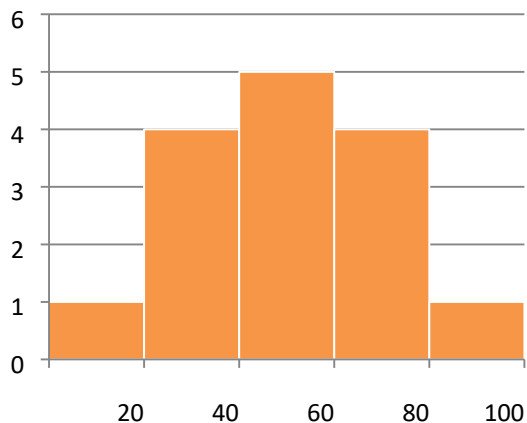


Exercices

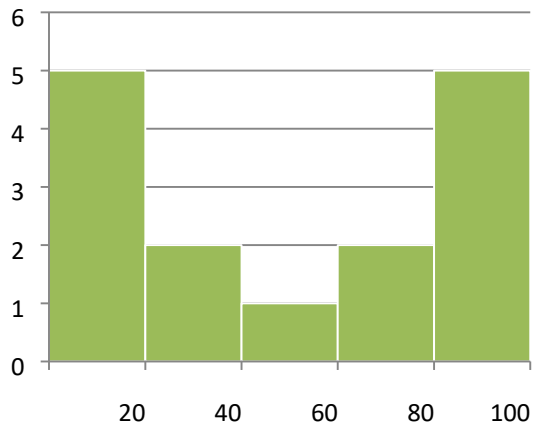
Exercice 1

Quatre études similaires ont été réalisées dans quatre pays différents et ont donné les histogrammes ci-dessous :

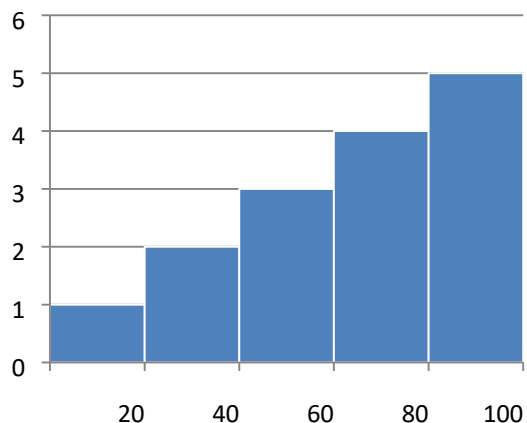
Étude A



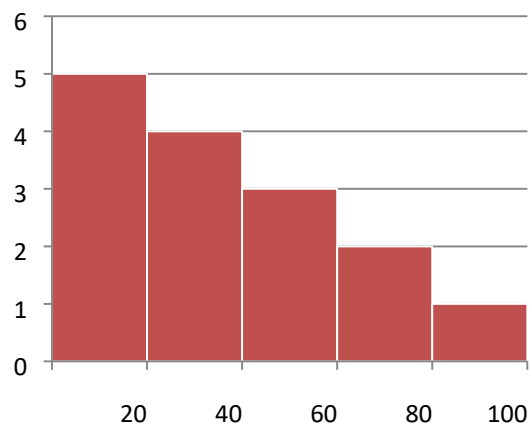
Étude B



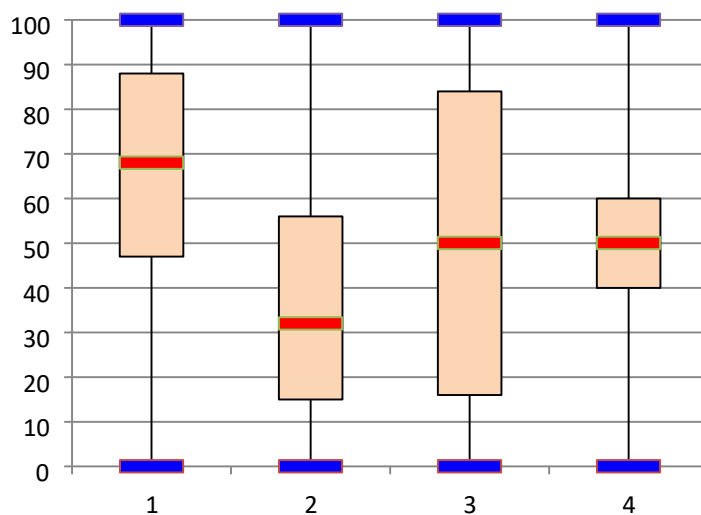
Étude C



Étude D



Associer ces quatre histogrammes aux quatre diagrammes en boîte à moustaches donnés ci-dessous. **Justifier** vos choix.



.....

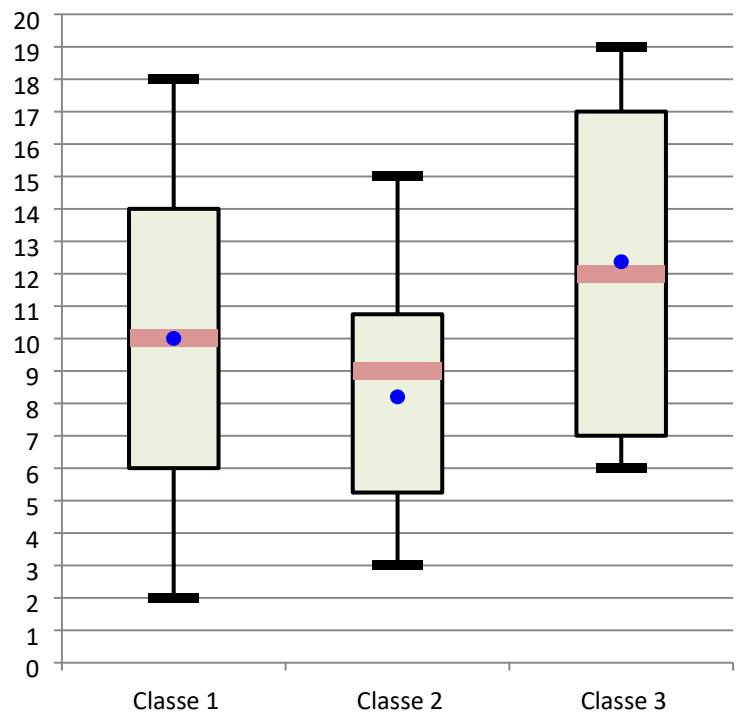
.....

.....

.....

Exercice 2

On donne 3 diagrammes en boîte à moustache correspondant aux résultats de 3 classes de première Bac Pro à un même devoir de mathématiques (la moyenne est figurée par un point).



1) Compléter le tableau suivant :

	Classe 1	Classe 2	Classe 3
Valeur minimale			
Q ₁			
Médiane			
Q ₃			
Valeur maximale			
moyenne			

À partir des diagrammes en boîtes à moustaches correspondant à chaque classe, donner le profil de chaque classe

.....

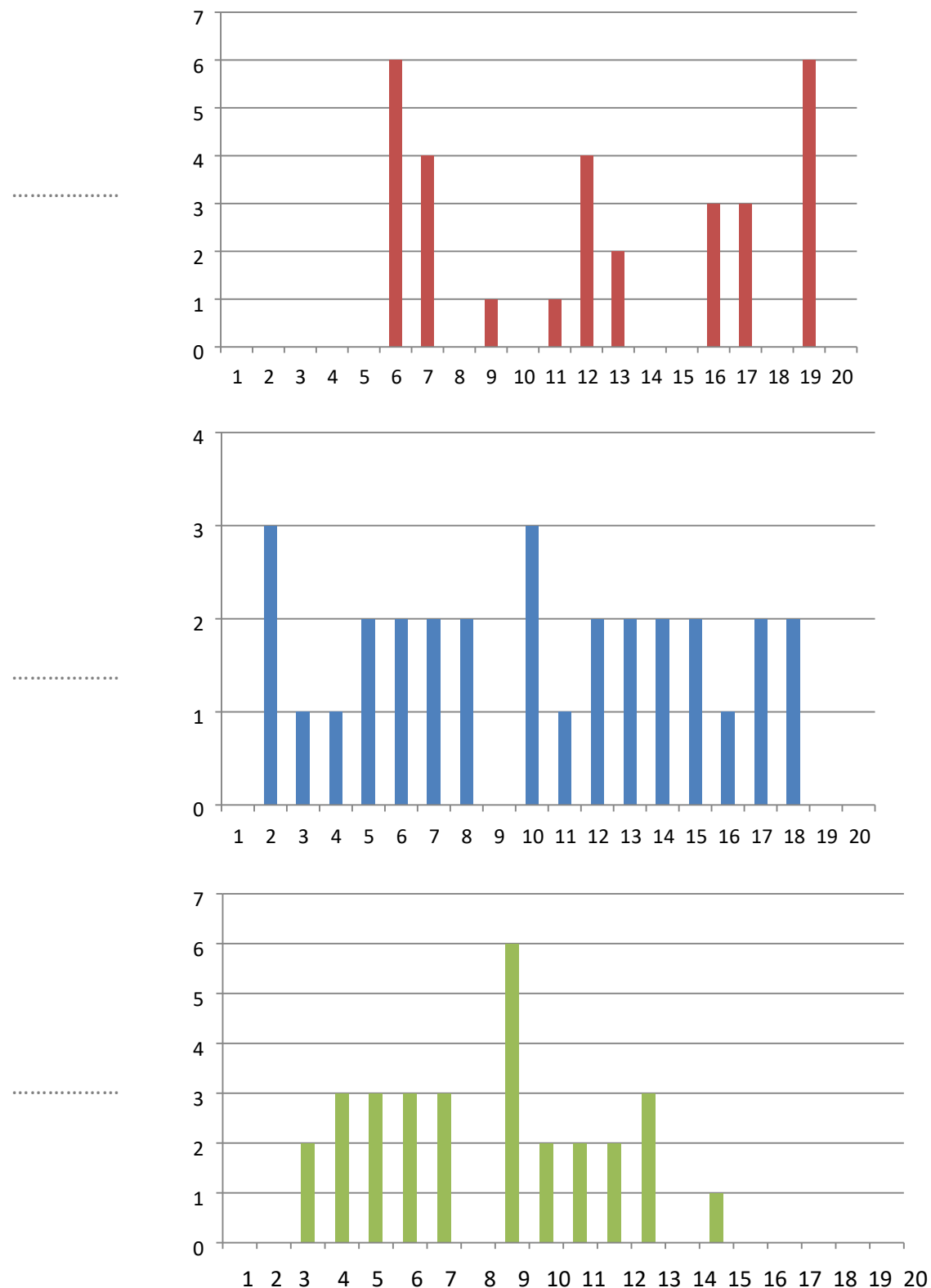
.....

.....

2) **Attribuer** le numéro de la classe en face de chaque diagramme en bâtons.

Classe

Diagramme à bâtons



5. Variance et Ecart type

1^{er} cas : la population est donnée par la liste de ses n éléments $x_1, x_2, \dots, x_1, \dots, x_n$.

2^e cas : la population est donnée par le tableau des effectifs n_i de chacune des p classes x_i .

3^e cas : la population est donnée par le tableau des effectifs n_i de chacune des p classes $[a_i, b_i[$ de centre $c_i = \frac{a_i + b_i}{2}$.

	Moyenne $\bar{x}$	Variance V
1 ^{er} cas	$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$	$V = \frac{x_1^2 + x_2^2 + \dots + x_n^2}{n} - \bar{x}^2$
2 ^e cas	$\bar{x} = \frac{n_1 x_1 + \dots + n_p x_p}{n}$	$V = \frac{n_1 x_1^2 + \dots + n_p x_p^2}{n} - \bar{x}^2$
3 ^e cas	$\bar{x} = \frac{n_1 c_1 + \dots + n_p c_p}{n}$	$V = \frac{n_1 c_1^2 + \dots + n_p c_p^2}{n} - \bar{x}^2$

L'écart-type de la série statistique est :

$$e = \sqrt{V}$$

L'écart type est la mesure la plus courante de la dispersion ou de la **répartition des données** sur la moyenne.

Interprétation

Une valeur d'écart type élevée indique que les données sont dispersées. D'une manière générale, pour une loi normale, environ 68 % des valeurs se situent dans un écart type de la moyenne, 95 % des valeurs se situent dans deux écarts types et 99,7 % des valeurs se situent

Méthode

Utilisation de la calculatrice La plupart du temps, on utilisera la calculatrice ^a pour déterminer l'écart-type d'une série.

On rentre les valeurs et les effectifs dans les listes de la calculatrice. L'écart-type est alors calculé et noté σX :

^a. voir par exemple sur le site

: <http://www.mathslycee.fr/2013-09-15-16-16-29/premiere-s/132-statistiques-et-probabilites/statistiques-descriptives/109-1s-sp1-stats-et-calculatrices>

```
1-Var Stats
x̄=83.66336634
Σx=16900
Σx²=1566600
Sx=27.56170002
σx=27.49339335
↓n=202
```

FIGURE 4 – écart-type avec une calculatrice TI

EXEMPLES

- Les 31 étudiants d'une formation professionnelle ont obtenu les notes suivantes à un contrôle de mathématiques.

Notes : x_i	7	8	9	10	11	12	13	14
Effectif : n_i	1	5	4	12	5	3	0	1

On est dans le 2° cas. Vérifier avec votre calculatrice que la valeur approchée arrondie à 10^{-1} de la moyenne est $\bar{x} \approx 9,9$ et que la valeur approchée arrondie à 10^{-1} de l'écart type est $\sigma \approx 1,5$.

- On effectue une enquête auprès de 120 grandes surfaces afin de déterminer le prix moyen d'une lampe solaire de 8 leds. On relève ces résultats dans le tableau suivant :

Prix de la lampe en euros	Nombre de grandes surfaces : n_i
[20, 25[	24
[25, 30[	32
[30, 35[	48
[35, 40[	12
[40, 45[	4

On fait l'approximation suivante : dans chaque classe, tous les éléments sont situés au centre de la classe.

On cherche le prix moyen $\bar{x}$ de la lampe solaire.

On est dans le 3° cas, avec $c_1 = 22,5$, $c_2 = 27,5$..., $c_5 = 42,5$. Vérifier avec votre calculatrice que $\bar{x} = 30$.

Exercice :

Les températures en (°C) relevées dans 15 villes différentes sont les suivantes :

25 28 26 28 30 32 25 27 28 30 31 32 26 31 33

Déterminer l'écart-type des températures.

.....

.....

.....

.....

.....