

Contrôle de Rattrapage de Statistique Descriptive  
SMA/S4. Le 29 juin 2019  
Durée 30'

Nom et Prénom :.....

Numéro Apogée :.....

Numéro de Table :.....

**Exercice 1 :**

Un fabricant d'ordinateurs a compté le nombre d'erreurs d'assemblage sur 396 d'entre eux et les résultats obtenus sont les suivants :

Nombre d'erreurs	Nombre d'ordinateurs	.....	.....	.....
0	101			
1	140			
2	92			
3	42			
4	18			
5	3			
Total				

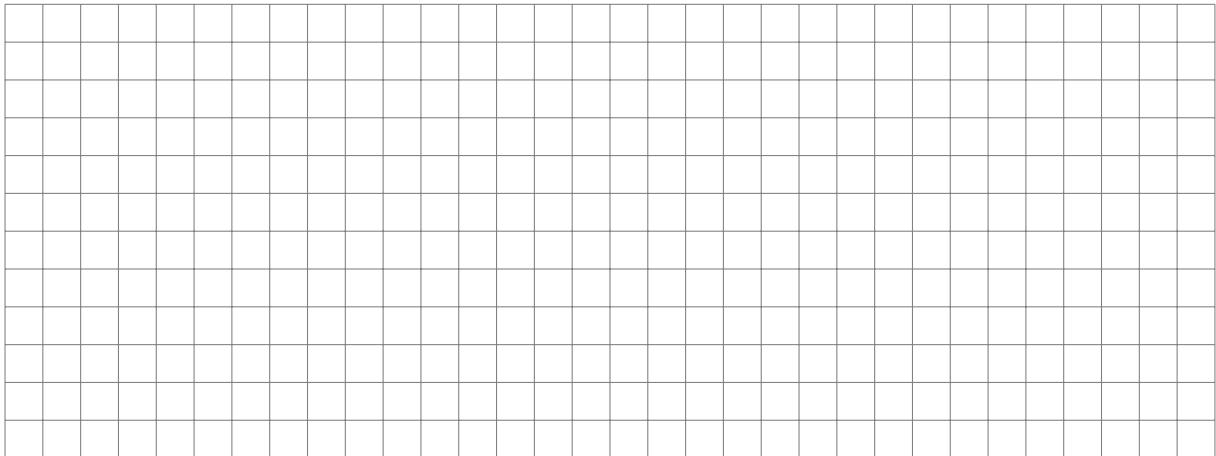
1) Quels sont la population et l'échantillon considérés ?

.....  
.....  
.....  
.....  
.....

2) Quelle est la variable étudiée et quelle est sa nature ?

.....  
.....  
.....  
.....

**3)** Représenter graphiquement ces données.



**4)** Calculer la moyenne, la variance et l'écart-type de cette distribution. Donner une interprétation à chacune de ces valeurs.

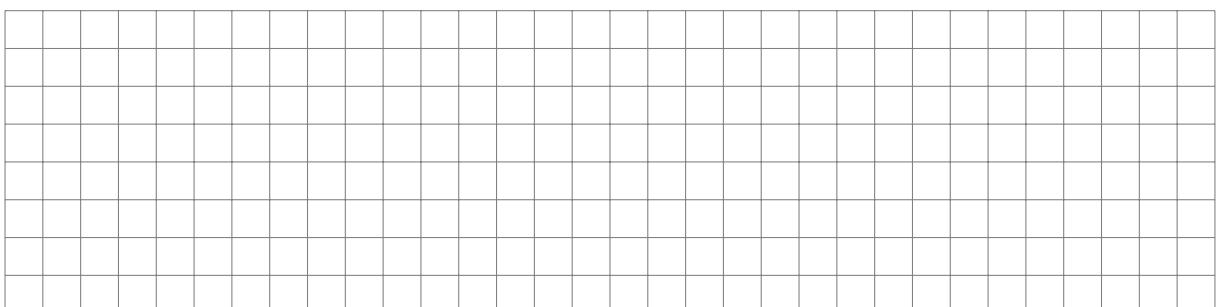
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

**5)** Déterminer les trois quartiles  $Q_1$ ,  $Q_2$  et  $Q_3$  en justifiant vos calculs. Dire ce que signifie chacun d'entre eux.

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

**6)** Tracer le diagramme en boîte correspondant à ces données.

.....  
.....  
.....



**7)** Existe-t-il des valeurs aberrantes ? Si oui, en donner une interprétation.

.....  
.....  
.....  
.....

**Exercice 2 :**

Une étude statistique a été faite, dans une certaine faculté, pour connaître l'impact de l'assiduité des étudiants sur leurs résultats aux contrôles.

On a choisi, au hasard, un échantillon de 333 étudiants de cette faculté et on a considéré deux variables  $X$  et  $Y$ .

La variable  $X$  représente l'“assiduité“ et prend deux modalités ”Assidu“ et ”Non assidu“.

La variable  $Y$  représente le ”niveau des résultats aux contrôles“ et prend trois modalités ”Faible“, ”Moyen“ et ”Bon“.

Les résultats obtenus sont résumés dans le tableau ci-dessous :

$X/Y$	Faible	Moyen	Bon	Total
Assidu	14	42	97	
Non assidu	103	51	26	
Total				

**1)** Quelle est la population étudiée ? Quel est l'échantillon considéré ? Préciser le type des variables utilisées.

.....  
.....  
.....  
.....  
.....

**2)** Déterminer la distribution de  $Y$  conditionnellement à  $X$ . Préciser comment on l'obtient.

.....  
.....  
.....

$X/Y$	Faible	Moyen	Bon	Total
Assidu				
Non assidu				

Que peut-on dire, à priori, du lien entre  $X$  et  $Y$ , juste en observant cette distribution pour l'échantillon ?

.....  
.....  
.....  
.....

Maintenant on voudrait étudier l'indépendance entre  $X$  et  $Y$  en utilisant le coefficient  $V$  de Cramér.

- 3)** Calculer les effectifs théoriques sous l'hypothèse d'indépendance. Justifier vos calculs.

.....  
.....  
.....

$X/Y$	Faible	Moyen	Bon	Total
Assidu				
Non assidu				
Total				

- 4)** Calculer les valeurs de  $q^2$  et de  $\phi^2$  puis en déduire celle du coefficient  $V$  de Cramér. Que peut-on dire de l'intensité du lien entre  $X$  (assiduité) et  $Y$  (niveau des résultats) pour l'échantillon.

**Exercice 1 :**

Une étude faite sur les exploitations agricoles dans une certaine région a consisté à relever les superficies (en hectares) de ces exploitations selon le nombre d'exploitants (propriétaires). Les résultats obtenus sont résumés dans le tableau suivant :

Superficies (hectars)	$n_i$	.....	.....	.....	.....	.....	.....	.....
[0; 5[	110							
[5; 10[	150							
[10; 15[	90							
[15; 20[	130							
[20; 40[	60							
[40 : 50[	50							
[50; 100[	40							
[100; 200[	20							
Total								

- 1)** Après avoir complété le tableau ci-dessus, en calculant les paramètres nécessaires, tracer l'histogramme des effectifs.



- 2)** Calculer la moyenne, la variance et l'écart-type de cette distribution.

---

---

---

---

---

- 3)** Cette distribution admet-elle un mode ? Interpréter.

---

---

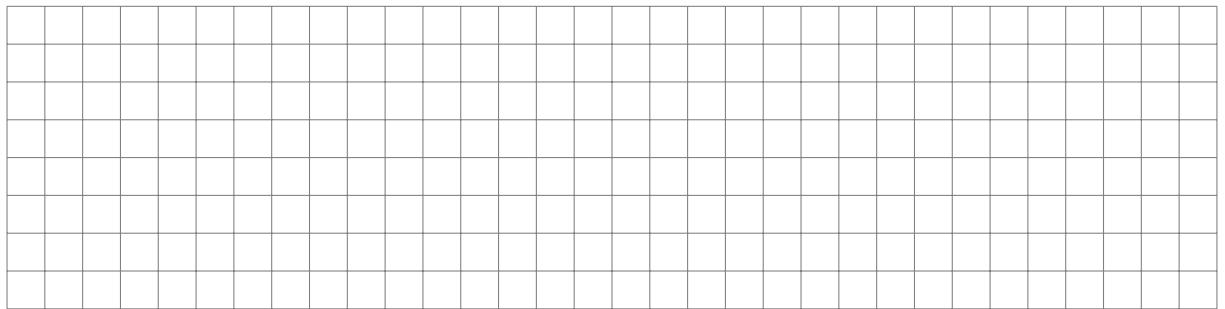
---

---

- 4)** Déterminer les trois quartiles  $Q_1$ ,  $Q_2$  et  $Q_3$ . Donner une interprétation à chacun de ces paramètres.

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

**5)** Tracer le diagramme en boîte associé à ces données



**6)** Existe-t-il des valeurs aberrantes ? Commenter.

.....  
.....  
.....  
.....  
.....

**Exercice 2 :**

Dans une étude sur la malnutrition dans un pays pauvre, on a calculé le poids moyen  $Y$  (en kg) par tranche d'âge  $X$  d'un échantillon de 2400 enfants répartis uniformément en 12 classes d'âge.

On a résumé les résultats dans le tableau suivant :

$x_i = \text{classe d'âge}$	1	2	3	4	5	6	7	8	9	10	11	12
$y_i = \text{poids moyen}$	3.5	3.5	3.3	4.4	4.4	4.2	5.1	5.3	5.5	5.5	6.2	5.7
$x_i^2$												
$y_i^2$												
$x_i y_i$												

- 1)** Représenter le nuage de points  $(x_i, y_i); i = 1, \dots, 12$ .



- 2)** Compléter le tableau ci-dessus puis déterminer l'équation de la droite qui ajuste le mieux le nuage, au sens des moindres carrés. Dessiner-la sur le même plan que le nuage.

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

- 3)** Calculer le coefficient de corrélation linéaire  $r_{XY}$  entre  $X$  et  $Y$ . Que peut-on dire de l'intensité du lien linéaire entre le poids et l'âge ?

.....  
.....  
.....  
.....  
.....

TOURNEZ LA PAGE SVP

4) Calculer les résidus  $(e_i)_{i=1,\dots,12}$ , puis la moyenne résiduelle. Commenter.

$x_i = \text{classe d'âge}$	1	2	3	4	5	6	7	8	9	10	11	12
$y_i = \text{poids moyen}$	3.5	3.5	3.3	4.4	4.4	4.2	5.1	5.3	5.5	5.5	6.2	5.7
$\hat{y}_i$												
$e_i$												

## V.S.CONTINUE + CALCUL DES QUARTILES + DIAGRAMME EN BOITE

### Exercice 1 :

La mesure de la taille  $X$ , en centimètres, de 30 sardines pêchées à Essouira a donné les résultats suivants :

22.0 15.6 16.4 20.5 17.4 15.9 21.8 14.9 18.7 23.7 18.6 19.0 21.5 22.3 25.1  
23.1 18.5 19.6 16.7 25.1 13.8 14.7 23.4 22.2 15.6 20.5 21.8 17.8 23.0 18.0

- a) Calculer la moyenne, la variance et l'écart-type de ces tailles. On donne  $\sum_{i=1}^{30} x_i^2 = 11\,801.82$

.....  
.....  
.....  
.....  
.....  
.....  
.....

- b) En considérant 4 classes d'amplitudes respectives 2, 4, 4 et 3 et dont la borne inférieure de la première classe est 13, remplir la 1<sup>ère</sup> colonne du tableau suivant :

Classes de tailles	.....	.....	.....	.....	.....	.....	.....
[.....; .....[							
[.....; .....[							
[.....; .....[							
[.....; .....[							

- c) Après avoir complété le tableau ci-dessus, dessiner l'histogramme des fréquences correspondant à ces données. Utiliser le quadrillage en fin d'épreuve.

- d) En utilisant les données groupées du tableau, calculer les nouvelles valeurs de la moyenne, la variance et l'écart-type. Comparer ces nouvelles valeurs à celles trouvées dans a) puis commenter.

.....  
.....  
.....  
.....  
.....

- e) Déterminer les trois quartiles puis dessiner le diagramme en boîte. Utiliser le quadrillage en fin d'épreuve.

.....  
.....  
.....  
.....  
.....  
.....

f) Calculer le quantile d'ordre 0.35.

**Exercice 2 :**

Dans le but de mieux comprendre l'effet de l'exposition quotidienne au soleil sur la production de tomates, on a effectué une expérience dans une serre du sud marocain ensoleillé la plupart des jours de l'année. On a planté 20 plantes de tomates, qu'on a disposées de sorte que :

- \* 4 plantes reçoivent 4 heures de soleil par jour,
- \* 4 plantes reçoivent 5 heures de soleil par jour,
- \* 4 plantes reçoivent 6 heures de soleil par jour,
- \* 4 plantes reçoivent 7 heures de soleil par jour.
- \* 4 plantes reçoivent 8 heures de soleil par jour.

Quatre mois plus tard, à la fin de l'expérience, on a pris note de la production totale de chaque plante. Les résultats, en kg, sont les suivants :

Nombre d'heures d'ensoleillement $X$	Production de tomates $Y$	$\bar{y}_i$
4 heures	2.54 2.92 3.65 3.57	
5 heures	3.17 2.72 4.53 4.48	
6 heures	4.13 4.64 4.55 3.93	
7 heures	5.20 4.63 4.44 5.08	
8 heures	5.43 3.99 5.76 4.86	

1) Compléter le tableau en calculant la moyenne de production de tomates pour chaque nombre d'heures d'ensoleillement.

2) Représenter le nuage de points  $\{(x_i, \bar{y}_i)_{i=1,\dots,5}\}$  dans un plan orthogonal. Utiliser le quadrillage en fin d'épreuve.

3) Peut-on deviner le signe et la valeur du coefficient de corrélation linéaire entre  $X$  et  $Y$  ?

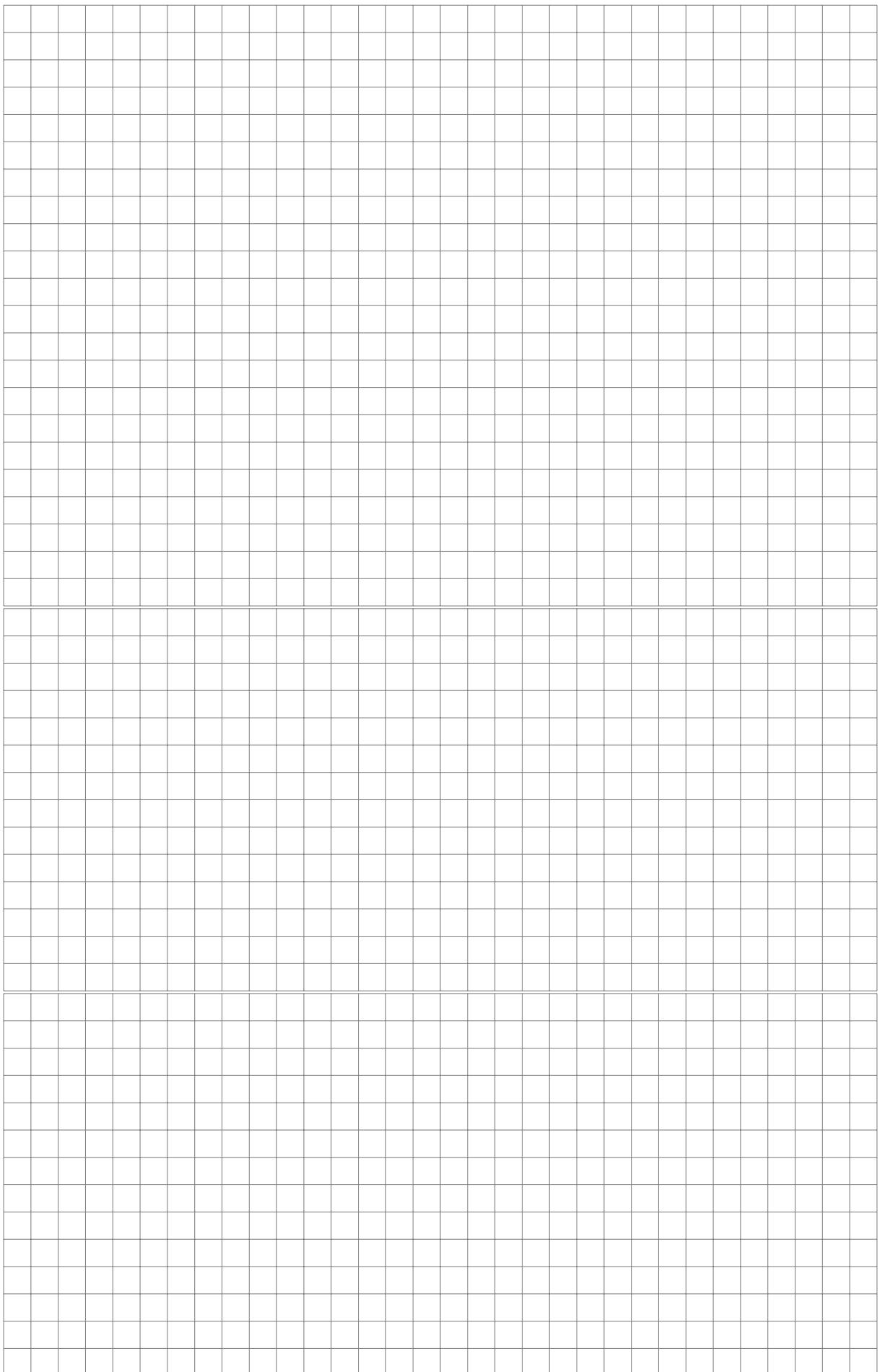
.....  
.....  
.....

4) Calculer ce coefficient de corrélation linéaire.

.....  
.....  
.....  
.....  
.....  
.....  
.....

5) Déterminer la droite de régression de  $Y$  en  $X$  puis la tracer sur le même plan où se trouve le nuage. Commenter.

.....



## Questions cours/

Etant donnée une série de nombres  $x_1, x_2, \dots, x_n$ .

1) A quoi correspond :

a)) le moment d'ordre un, pour ces nombres ?

.....  
.....  
.....  
.....

b)) le moment centré d'ordre un ?

.....  
.....  
.....  
.....

c)) le moment centré d'ordre deux ?

.....  
.....  
.....  
.....

2) Que mesure la variance de cette série ?

.....  
.....  
.....  
.....  
.....  
.....

## Exercice/

Une étude sur les prix de tablettes électroniques (en dirhams) a donné les résultats suivants :

Prix (dh)	Fréquences cumulées	.....	.....	.....	.....
[800 ; 1000[	0.08	.....	.....	.....	.....
[1000 ; 1400[	0.18	.....	.....	.....	.....
[1400 ; 1600[	0.34	.....	.....	.....	.....
[1600 ; y[	0.64	.....	.....	.....	.....
[y ; 2400[	0.73	.....	.....	.....	.....
[2400 ; x[	1	.....	.....	.....	.....

1) Certaines données sont manquantes. Calculer la borne manquante  $x$  sachant que l'étendue de la série est égale à 3200 dirhams.

.....  
.....  
.....  
.....  
.....  
.....

**2)** Calculer la borne manquante  $y$  dans les deux cas suivants :

a) Le prix moyen est de 1995 dirhams.

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

b) Le prix médian est de 1920 dirhams.

.....  
.....  
.....  
.....  
.....  
.....  
.....

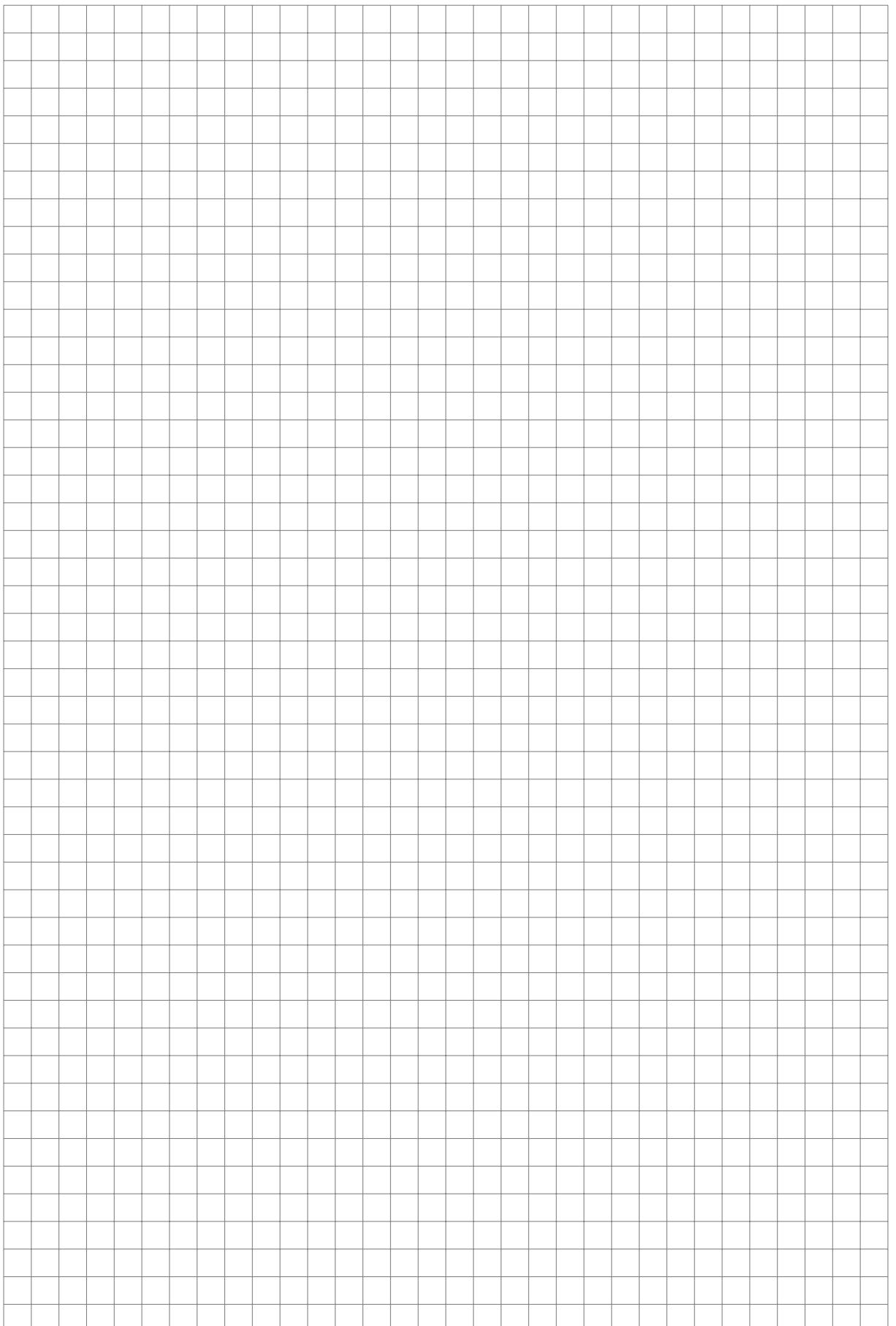
**On aura besoin, éventuellement, de calculer les fréquences relatives à chaque classe.**

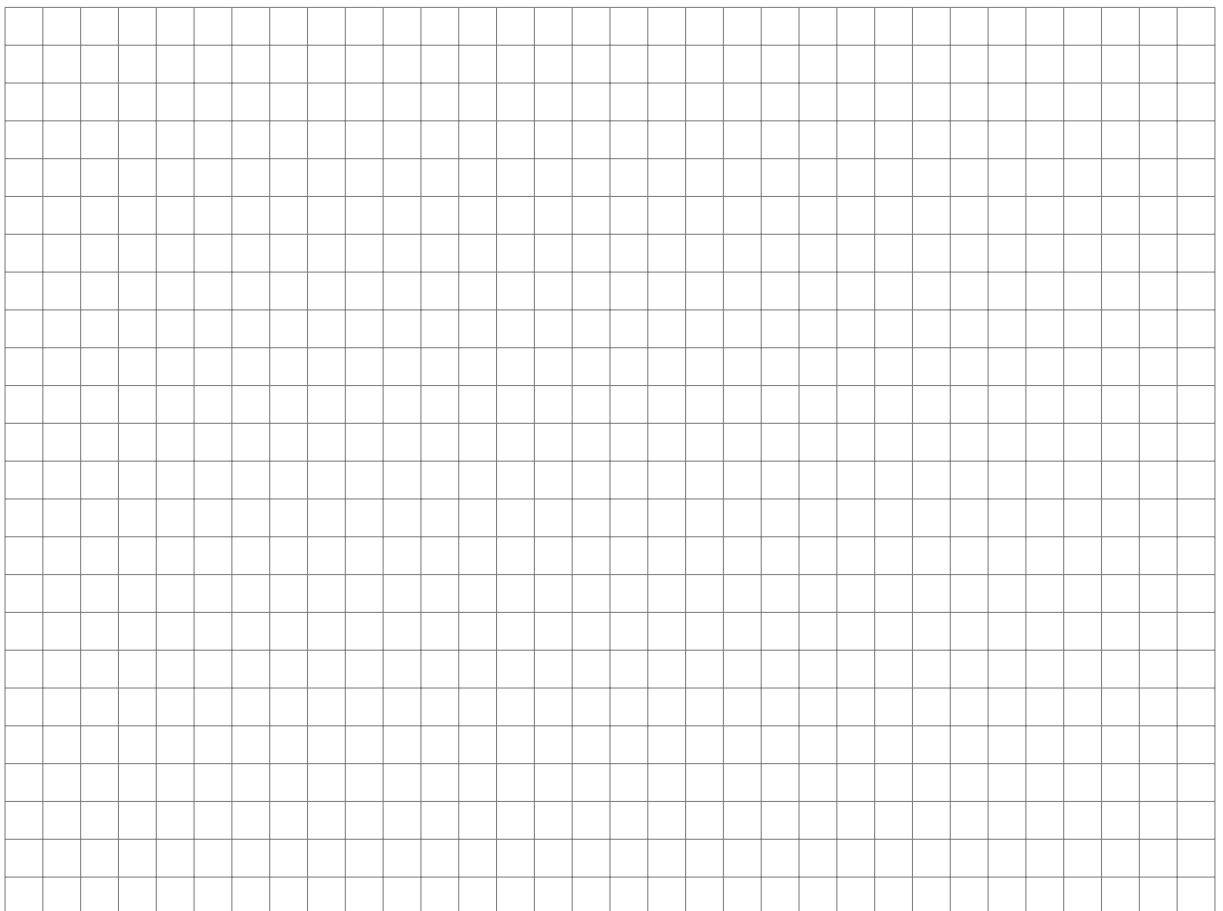
**3)** On suppose dans ce qui suit que  $y = 2000$  dirhams.

a) Après avoir complété le tableau ci-dessus, donner une représentation graphique de ces prix. Utiliser le quadrillage en fin d'épreuve.

b) Recalculer et interpréter les prix moyen et médian. Que peut-on déduire de la comparaison de ces deux valeurs ?

.....  
.....  
.....  
.....  
.....  
.....





## Exercice

Donner 5 entiers naturels tous distincts dont :

- a) La moyenne vaut 5 et la médiane vaut 5.
  - b) La moyenne vaut 5 et la variance vaut 6.
  - c) La moyenne vaut 9, l'étendue vaut 20, la médiane vaut 5.

## Problème

On a relevé les longueurs  $x_k$  ( $k = 1, \dots, 25$ ) suivantes, en mm, de 25 poissons :

214; 220; 220; 225; 168; 178; 176; 176; 182; 200; 185; 175; 189;  
164; 195; 210; 181; 175; 170; 204; 190; 194; 190; 223; 180

- 1) Quels sont la population et l'échantillon concernés par l'étude ?

.....  
.....  
.....  
.....

- 2)** Quel est le caractère étudié ? Donner sa nature.

---

---

---

---

---

- 3) Avec ces données brutes, calculer la médiane, la moyenne et la variance des longueurs des poissons. ( $\sum_{k=1}^{25} x_k = 923664$ )

- 4) Ranger les données précédentes, en utilisant les classes proposées dans le tableau ci-dessous.

On complètera les colonnes du tableau par les valeurs nécessaires pour le calcul des différents paramètres demandés par la suite.

Classes	Amplitudes	.....	.....	.....	.....	.....
[164; 176[	.....	.....	.....	.....	.....	.....
[176; 186[						
[186; 196[						
[196; 211[						
[211; 226[						

Tracer l'histogramme correspondant à ce tableau. Utiliser le quadrillage en fin d'épreuve.

- 5) Calculer les nouvelles valeurs de la médiane, de la moyenne et de la variance en utilisant les données groupées.

Puis comparer-les à celles obtenues ci-dessus et commenter.

