Dans le chapitre 10 vous avez appris à être plus autonome dans la lecture, l'interprétation et la réalisation de tableaux et de graphiques que l'on rencontre dans l'étude de situations issues de la vie économique et sociale.

Pour rendre compte d'une série statistique, on la caractérise par des nombres significatifs. On définit des paramètres qui sont de deux sortes : les paramètres de position (mode, moyenne, médiane) et les paramètres de dispersion (étendue, écart moyen, variance et écart type).

A la fin de ce chapitre, vous devrez connaître tous ces paramètres.

1 Moyenne.

La moyenne arithmétique simple de p nombres est obtenue en calculant la somme de ces p nombres, puis en divisant cette somme par p.

$$M = \frac{x_1 + x_2 + ... + x_p}{p} = \frac{\text{Somme des nombres}}{p} = \frac{1}{p} \times \sum_{i=1}^{p} x_i$$

Le symbole sigma signifie qu'il faut additionner tous les termes x_i avec i prenant les valeurs de 1 jusqu'à p.

Soit une série statistique prenant les valeurs x_1 , $x_{2,}$ …, x_p avec les effectifs n_1, n_2, … , $n_{p,}$

la moyenne de la série est : $\overline{x} = \frac{n_1 x_1 + n_2 x_2 + ... + n_p x_p}{n_1 + n_2 + ... + n_p} = \frac{\text{Somme des nombres affectés de leurs coefficients}}{n_1 + n_2 + ... + n_p}$

$$\overline{x} = \frac{\sum_{i=1}^{p} n_i x_i}{\sum_{i=1}^{p} n_i}$$

Dans le cas d'une série dont les valeurs ont été regroupées en classes, les valeurs x_i sont les centres des classes.

Lorsque dans une série statistique, il y a une valeur exceptionnelle, il peut être intéressant de calculer la moyenne de cette série privée de la valeur exceptionnelle. On dit alors que l'on a calculé une moyenne élaguée.

Soit P une population partagée en deux sous populations disjointes d'effectifs p et q.

Autrement dit on répartit n nombres en deux sous-groupes disjoints l'un contenant p éléments et l'autre contenant q éléments.

Si $\bar{x}$ est la moyenne des p nombres du premier sous-groupe et si $\bar{y}$ est la moyenne des q nombres du second sous-groupe, alors la moyenne des n nombres est égale à $M = \frac{p\bar{x}+q\bar{y}}{n} = \frac{p}{n}\bar{x} + \frac{q}{n}\bar{y}$.

E1 Savoir calculer des moyennes.

N ° 1

Les notes de maths de Tom sont 15 ; 18 ; 16 ; 12 ; 8 et 9. Calculer sa moyenne.

N ° 2

On considère la série statistique donnée par le tableau suivant :

Valeur du caractère	8	9	10	11	12	18
Effectif	7	5	8	9	4	1

Calculer la moyenne de cette série.

N ° 3

Dans la classe de première STG, il y a 31 élèves dont 19 filles et 12 garçons.

La moyenne de maths des filles du 2T est égale à 14,27.

La moyenne de maths de la classe du 2T est égale à 13,69.

Calculer la moyenne de maths des garçons du second trimestre.

E2 Activité pour comprendre l'intérêt de l'écart type.

N ° 4 Voici le relevé des notes d'un élève de première STG dans les matières générales.

Français	11	12	11	10	09
Mathématiques	16	08	10	06,5	13,5
Anglais	15	15	15	05	05
Histoire géographie	11	11,5	11	10	11

1. Niveau général de l'élève

a) Sans calcul, peut on dire dans quelle matière l'élève a la meilleure moyenne.
b) Calculer la moyenne de l'élève dans chaque matière.

2. Solidité des résultats

a) Sur la calculatrice, en mode statistiques, rentrer les 4 séries de notes dans 4 listes différentes.

Pour chacune des listes, utiliser l'instruction 1VAR du menu CALC et lire l'écart type.

(Il faut rentrer dans STATS LIST 1 les notes de français puis CALC set 1 var X : List 1 ; 1 var F : 1

QUIT CALC 1 var et la moyenne est donnée par la lettre $\bar{x}$ et l'écart type est donné par la notation x σ n .)

b) Recopier les quatre écarts types trouvés.
c) Pour chaque matière, observer l'homogénéité des notes. Quelle indication donne l'écart type correspondant ?

2 Ecart type.

La variance V est la moyenne des carrés des écarts entre chaque valeur x_i et la moyenne $\overline{x}$.

La variance permet de mesurer la dispersion des valeurs autour de la moyenne.

La variance V est donnée par les formules ci-dessous :

$V = \frac{1}{N}\sum_{i=1}^{n} n_i(x_i - \overline{x})^2$ ou bien $V = \frac{1}{N}\sum n_i x_i^2 - \overline{x}^2$

L'écart-type s d'une série statistique est la racine carrée de la variance.

Les anglo-saxons parlent de " standard deviation " d'où la lettre s pour désigner l'écart type.

L'écart-type est exprimé dans la même unité que la variable.

L'écart type permet de chiffrer la dispersion d'une série statistique autour de la moyenne.

Plus il est fort, plus les valeurs sont dispersées.

L'écart type varie fortement lorsqu'on modifie des valeurs extrêmes de la série.

On peut remplir un tableau pour présenter les calculs de moyenne et de variance.

Avec une calculatrice il faut rentrer dans stats list 1 les valeurs du caractère ou les centres des classes ;

dans list 2 les effectifs correspondants; puis calc set 1 var X : list 1 ; 1 var F : list 2 quit calc 1 var…

E3 Savoir calculer un écart type avec la calculatrice.

N ° 5 Monsieur et madame Melon jouent régulièrement au scrabble avec leur voisine.
Voici leurs scores des quinze dernières parties.

Madame	166	220	152	235	145	175	238	82	183	217	169	205	180	178	250
Monsieur	197	240	271	175	254	163	181	200	160	173	169	230	179	189	199

1. Calculer la moyenne et l'écart type de chacun.
2. Lequel des deux semble avoir le meilleur niveau ?
3. Lequel des deux semble le plus régulier ?

N ° 6
Au tennis, l'entraîneur d'un joueur de haut niveau a relevé la vitesse de son premier service au cours d'un match.

1. Calculer la vitesse moyenne du premier service.
2. Calculer l'écart type de la série.

Vitesse en km/h	[132 ; 140 [	[140 ; 148 [	[148 ; 156 [	[156 ; 164 [
Nombre de services	4	7	13	24

Vitesse en km/h	[164 ; 172 [	[172 ; 180 [	[180 ; 188 [	[188 ; 196 [
Nombre de services	53	30	9	5

3 Quantiles.

Dans ce paragraphe, on considère la liste des valeurs d'une série statistique à caractère quantitatif, triées dans l'ordre croissant, chacune apparaissant le nombre de fois indiqué par son effectif.

Le principe est de découper la liste ordonnée ainsi constituée, en listes ayant toutes le même effectif.

Médiane

La médiane M_e de la série est un nombre qui découpe la liste ordonnée en deux listes ayant sensiblement le même effectif.

Autrement dit : la médiane est la valeur centrale de la série si n est impair ou bien la demi somme des deux valeurs centrales si n est pair.

Dans le cas d'un caractère quantitatif continu (en supposant une répartition régulière à l'intérieur de chaque classe), une valeur approchée de la médiane est l'abscisse du point d'ordonnée $\frac{n}{2}$ du diagramme des effectifs cumulés croissants.

On parle alors de classe médiane.

Schéma : voir feuille annexe.

Quartiles

Les quartiles Q_1, Q_2, Q_3 de la série sont trois valeurs de la série qui découpent la liste ordonnée en quatre listes ayant sensiblement le même effectif.

Autrement dit : dans le cas de valeurs non regroupées en classe, si $\frac{N}{4}$ est un entier n, alors le premier quartile Q_1 est le terme qui dans cette liste occupe le rang n et le troisième quartile Q_3 est le terme de rang 3n.

Si $\frac{N}{4}$ n'est pas un entier , alors le premier quartile Q_1 est le terme qui dans cette liste occupe le rang strictement supérieur à $\frac{N}{4}$ et le troisième quartile Q_3 est le terme de rang strictement supérieur à $\frac{3N}{4}$.

Dans le cas de valeurs regroupées en classes, Q_1 est la valeur correspondant à la fréquence cumulée croissante égale à 0,25 et Q_3 est la valeur correspondant à la fréquence cumulée croissante égale à 0,75.

Schéma : voir feuille annexe.

Déciles

Les déciles D_1, D_2, D_3, …, D_9 de la série sont neuf valeurs de la série qui découpent la liste ordonnée en dix listes ayant sensiblement le même effectif.

Schéma : voir feuille annexe.

Exemple :

En colonie de vacances, l'animateur qui organise les olympiades compte 25 inscrits pour courir le 100 mètres.
Il décide de constituer quatre groupes d'à peu près même effectif qui prendront le départ successivement des plus petits aux plus grands. Voici les tailles des candidats en cm :

138, 131, 128, 150, 144, 121, 154, 120, 127, 149, 125, 127, 133, 149, 140, 129, 153, 160, 150, 142, 144, 139, 153, 145, 161.

Déterminons la médiane et les quartiles Q_1 et Q_3.

E4 Savoir déterminer des quantiles.

N ° 7

Les résultats d'une enquête sur le prix de la baguette de pain dans les 30 boulangeries d'une ville sont regroupées dans le tableau suivant :

prix en centimes d'euros	68	70	72	74	75	76	78	80	85	90	92
nombre de boulangeries	1	1	2	4	7	3	3	4	2	2	1

1. Dresser le tableau des effectifs cumulés croissants.
2. En déduire la médiane, les premier et troisième quartiles et les premier et neuvième déciles.

N ° 8

Dans un magasin de chaussures pour femmes, on a relevé les pointures des clientes durant une semaine.

pointures	35	36	37	38	39	40	41	42
effectifs	8	15	23	30	37	22	9	6

Déterminer la médiane, les premier et troisième quartiles et les premier et neuvième déciles.

4 Intervalle interquartile, intervalle interdécile.

L'écart interquartile est égal au nombre $Q_3 - Q_1$.

L'intervalle interquartile est [Q_1 ; Q_3].

Cet intervalle est une mesure de dispersion.

L'écart interdécile est égal au nombre $D_9 - D_1$.

L'intervalle interdécile est [D_1 ; D_9].

Cet intervalle est une mesure de dispersion.

Le diagramme en boîte d'une série est un diagramme regroupant la médiane, le premier et le troisième quartile.

On le nomme aussi " box - plot ".

Il existe aussi le diagramme en boîte à moustaches obtenu en plaçant le premier et le neuvième décile.

Les diagrammes en boîtes résument graphiquement une série.

Ce diagramme permet de visualiser la dispersion d'une série statistique.

Ils permettent de comparer plusieurs séries statistiques.

On construit un diagramme en boîte de la façon suivante :

Les valeurs sont disposées sur une droite parallèle à l'axe des abscisses (ou bien l'axe des ordonnées).

On place sur cet axe le minimum, le premier quartile, la médiane, le troisième quartile et le maximum de la série.

On construit alors un rectangle (la boîte) parallèlement à l'axe dont la longueur est l'interquartile et dont la largeur est arbitraire.

Exemple : Déterminer l'intervalle interquartile de l'exercice n ° 8.

Déterminer l'écart interquartile.

Déterminer l'intervalle interdécile de l'exercice n ° 8.

Résumer la série par un diagramme en boîte.

E5 Savoir déterminer l'intervalle interquartile d'une série.

N ° 9

Un professeur compare les notes obtenues à un test noté sur 10 par les deux groupes d'une classe.

groupe 1	8	3	7	2	5	7	9	6	8	3	3	8	6	5
groupe 2	6	7	3	5	6	6	8	4	7	8	6	7	5	6

1. a. Pour chaque groupe, déterminer la médiane, le premier et le troisième quartile.
   b. Pour chaque groupe, déterminer l'intervalle interquartile.
   c. Pour chaque groupe, en utilisant un même axe, réaliser les deux diagrammes en boîte correspondants.

2. Les utiliser pour comparer les deux groupes :
   a. concernant le niveau
   b. concernant l'homogénéité.

N ° 10

On donne dans le tableau ci dessous, le récapitulatif des distances parcourues chaque jour pendant un mois par un livreur de fleurs.

distances en km	[0 ; 10 [	[10 ; 20 [	[20 ; 30 [	[30 ; 40 [	[40 ; 50 [
nombre de jours	1	1	3	2	4

distances en km	[50 ; 60 [	[60 ; 70 [	[70 ; 80 [	[80 ; 90 [	[90 ; 100 [
nombre de jours	5	8	3	1	2

1. a. Dresser le tableau des effectifs cumulés croissants.
   b. Tracer le polygone des effectifs cumulés croissants.

2. a. Déterminer graphiquement la médiane, les premier et troisième quartiles et les premiers et neuvièmes déciles.
   b. Résumer ces résultats à l'aide d'un diagramme en boîte en utilisant l'axe des abscisses précédent.

E6 Exercice type bac.

N ° 11

Une machine fabrique des fers cylindriques pour le béton armé de diamètre théorique 25 mm. On contrôle le fonctionnement de la machine en prélevant un échantillon de 100 pièces au hasard dans la fabrication.
Les mesures des diamètres ont donné les résultats suivants :

classes par diamètre	[24 ; 24, 2 [	[24,2 ; 24,4 [	[24,4 ; 24,6 [	[24,6 ; 24,8 [	[24,8 ; 25 [
effectifs	0	5	13	24	19

classes par diamètre	[25 ; 25,2 [	[25,2 ; 25,4 [	[25,4 ; 25,6 [	[25,6 ; 25,8 [	[25,8 ; 26 [
effectifs	14	10	8	5	2

1. Réaliser un tableau en y mettant les centres des classes et les effectifs.

2. A l'aide de la calculatrice et en rappelant les formules utilisées, déterminer la moyenne $\overline{x}$ et l'écart type σ de la série statistique arrondis à 0,001. Interpréter les résultats trouvés.

3. La production de la machine est jugée bonne si la série des mesures de l'échantillon remplit les trois conditions suivantes :
a la moyenne appartient à l'intervalle [24,9 ; 25,1]
b l'écart type est strictement inférieur à 0,4
c 90 % au moins de l'effectif figure dans l'intervalle [$\overline{x} - 2\sigma$; $\overline{x} + 2\sigma$].
Vérifier les deux premiers critères.
4. Réaliser un tableau avec les classes par diamètres, les effectifs et les effectifs cumulés croissants.
5. Tracer le polygone des effectifs cumulés croissants en respectant les consignes suivantes :
sur l'axe des abscisses, on écrira 24 à 1 cm de l'unité et on prendra ensuite 1 cm pour 0,2 mm.
sur l'axe des ordonnées ; on choisira 1 cm pour 10 machines.
6. Déterminer le nombre de fers cylindriques de diamètres inférieur ou égal à 24,2 mm puis à 25,7 mm.
En déduire le nombre de fers cylindriques compris dans l'intervalle [$\overline{x} - 2\sigma$; $\overline{x} + 2\sigma$].
Calculer le pourcentage de fers cylindriques compris dans l'intervalle [$\overline{x} - 2\sigma$; $\overline{x} + 2\sigma$].
Vérifier le critère c.
7. La production de la machine est -elle bonne ?