STATISTIQUES

1 Généralités

1.1 C'est quoi ?

La statistique (par opposition à *une* statistique) est l'activité qui consiste à recueillir, traiter et interpréter un ensemble de données d'informations. Le traitement des données consiste à produire *des* statistiques (au pluriel).

Le but de la statistique est d'extraire *des informations pertinentes* d'une liste de nombres *difficile à interpréter* par une simple lecture.

1.2 Le vocabulaire

Population

C'est l'ensemble sur lequel porte l'étude. Qu'il s'agisse d'élèves dont on regarde les notes ou qu'il s'agisse de vis produites par un machine dont on regarde la taille, on parlera de la population (les élèves ou les vis).

Échantillon

C'est une partie de la population. Lorsque la population est trop grande, on étudie uniquement un échantillon de la population et on peut parfois généraliser les observations faites sur l'échantillon à la population entière. Ceci n'est possible que si l'échantillon est représentatif de la population entière. C'est le principe des sondages d'opinion. Les sondages se pratiquent aussi dans l'industrie : on extrait une quantité des vis produites et on vérifie si les dimensions sont proches de celles attendues. Si c'est le cas, on suppose que la machine fonctionne bien pour toutes les vis produites, sinon on cherche le problème dans la chaîne de production.

Individus

Ce sont les éléments qui composent la population. Avec l'exemple ci-dessus, un individu sera un élève ou une vis.

Caractère étudié, caractère observé

C'est l'aspect que l'on observe sur les individus. Ce pourra être la note en mathématiques à un devoir des élèves d'une classe, ou bien le diamètre des vis produites par une machine.

Série statistique

C'est l'ensemble des valeurs collectées. Par exemple la liste des notes obtenues par les élèves, ou la liste des dimensions des vis.

Série statistique quantitative

Si la série est une liste de nombres, ou, plus précisement, si le caractère observé est *mesurable*, on dira qu'il s'agît d'une série statistique *quantitative* (quantité = nombre).

Série statistique quantitative discrète

De plus, si le caractère étudié ne peut prendre qu'un nombre limité de valeurs, par exemple, le nombre d'enfants d'une famille (qui ne peut n'être que 0, 1, 2, 3, etc. mais pas 2,136), on dira qu'il s'agît d'une série statistique quantitative *discrète*.

Série statistique quantitative continue

Si le caractère étudié peut prendre un nombre illimité de valeurs (en général dans un intervalle), par exemple, la taille (qui peut être n'importe quelle valeur entre 0 m et … 3 m ?), on dira qu'il s'agît dans série statistique quantitative *continue*. Même si dans la pratique on limite le nombre de décimales.

Série statistique qualitative

La série peut être *qualitative* quand le caractère observé n'est pas mesurable. Une série de couleurs de cheveux, de mois de naissances, de sports pratiqués, etc. est une série qualitative.

1.3 Traitement des données

Classe

Quand il y a beaucoup de valeurs, on les regroupe par *classe* (intervalles), particulièrement dans le cas des séries statistiques continues. Par exemple si on étudie les tailles d'une population donnée importante, on pourra les regrouper les tailles qui vont de 1 m à 1,05 m ensemble, celles qui vont de 1,05 m à 1,10 m ensemble, etc.

Effectif, effectifs cumulés

Pour une valeur prise par le caractère étudié (pointure, couleur de cheveux, par exemple), c'est le nombre d'individus de la population ayant cette valeur. Les effectifs cumulés croissants s'obtiennent en ajoutant au fur et à mesure les effectifs des valeurs précédentes, mais nous verrons cela dans le détail plus tard.

Fréquence, fréquences cumulées

C'est le quotient de l'effectif d'une valeur prise par le caractère étudié par l'effectif total.

$$\text{Fréquence d'un caractère} = \frac{\text{effectif de ce caractère}}{\text{effectif total de la population}}$$

Les fréquences cumulées croissantes s'obtiennent en ajoutant au fur et à mesure les fréquences des valeurs précédentes. Là aussi nous le verrons plus tard.

Les fréquences s'expriment très souvent en pourcentage, mais ce n'est pas systématique.

1.4 Étude d'une série statistique

Nous allons nous intéresser essentiellement aux *séries statistiques quantitatives*.

Les séries étudiées en statistique sont parfois composées de nombreux éléments qu'il est difficile de manipuler tous à la fois. Aussi résume-t-on la série par plusieurs valeurs qui permettent de « faire parler » les nombreux éléments.

1.4.1 Mesures centrales

Les mesures centrales visent à *résumer toute la série en un seul nombre*, qu'on espère le plus représentatif possible de la série statistique.

Chacun de ces résumés possède des défauts et des qualités que nous verrons plus tard.

Mode, classe modale

Définition 1
Le mode *d'une série est la valeur la plus fréquente.*

Lorsque que les valeurs sont regroupées en classes, et que l'amplitude de chaque classe est la même, *la classe modale* est la classe qui a l'effectif le plus grand. Il y a parfois plusieurs modes, si deux valeurs de la série sont les plus fréquentes à égalité.

Le mode ou la classe modale n'ont vraiment d'intérêt que si leur effectif est nettement supérieur aux autres effectifs.

Il pourra intéresser le publicitaire.

Médiane : m

Définition 2
La médiane *est une valeur qui partage la population en deux parties de telle sorte que :*
- la moitié au moins *des données prennent une valeur* inférieure ou égale *à la médiane ;*
- la moitié au moins *des individus prennent une valeur* supérieure ou égale *à la médiane.*

Plusieurs valeurs peuvent parfois convenir pour la médiane, on parle alors d'*une* médiane au lieu de *la* médiane.

C'est le cas s'il y a un *nombre pair de données*. Par exemple, pour la série $\{6,7,8,10,11,14\}$, on peut choisir 8,5 car il y a au moins la moitié des données inférieures à 8,5 et au moins la moitié des données supérieures à ce nombre. Mais c'est le cas aussi pour 9, pour 10 et, plus généralement pour tous les nombres compris dans l'intervalle $[8;10]$.

Alors lequel choisir pour la médiane ? À vrai dire, le statisticien s'en moque, car, quand le nombre de données dans la série est très grand, par exemple 500, la 250^{e} donnée et la 251^{e} sont en général très proches et donnent toutes les deux un ordre de grandeur satisfaisant de la « donnée du milieu » et lorsque le nombre de données est petit, il n'y a pas beaucoup d'intérêt à faire des statistiques.

Mais en Seconde on travaillera sur des séries assez réduites et, puisqu'on attend un seul nombre comme médiane, on conviendra alors, conformément au programme, que la médiane est la moyenne des deux valeurs centrales. Ici $\frac{8+10}{2}=9$.

S'il y a un *nombre impair de données*, on n'a pas le choix. Par exemple pour la série $\{6,7,8,10,11,14,15\}$, un seul nombre peut convenir : 10. En effet 9, par exemple, ne convient pas puisque s'il y a au moins la moitié des valeurs de la série supérieures à 9, il n'y a par contre pas la moitié des valeurs de la série inférieures à 9.

Résumons-nous :

Propriété 1
Lorsque les valeurs d'une série statistique quantitative sont rangées dans l'ordre croissant et que :
- *le* ***nombre de valeurs*** n ***est impair****, la médiane est* obligatoirement *la* $\dfrac{n+1}{2}^{e}$ *valeur de la série.*
- *le* ***nombre de valeurs*** n ***est pair****, tout nombre compris entre la* $\dfrac{n}{2}^{e}$ *valeur et la valeur suivante est une médiane mais on convient de prendre la moyenne des deux.*

Lorsqu'on ne dispose que des fréquences, la médiane est la valeur qui correspond à la fréquence cumulée de 0,5.

La médiane a l'avantage de ne pas être sensible aux valeurs extrêmes.

Elle intéressera plutôt le sociologue.

Moyenne, moyenne élaguée : $\overline{x}$

Définition 3
La moyenne *d'une série statistique quantitative est la valeur telle que* la somme de toutes les valeurs de la série est la même si on remplace chaque valeur par la moyenne.

Par exemple la série $\{6,7,8,10,11,14,14\}$ a pour moyenne $\overline{x}=10$ car $6+7+8+10+11+14+14=70=10+10+10+10+10+10+10$.

Propriété 2
On trouve donc la moyenne avec la formule suivante :

$$\overline{x}=\frac{\textit{somme de toutes les valeurs}}{\textit{nombre de valeurs}}$$

La moyenne a le défaut d'être très sensible aux valeurs extrêmes. Par exemple la série $\{6,7,8,20\}$ a pour moyenne $\overline{x}=10,5$ qui n'est pas vraiment représentatif de la série.

C'est pourquoi, lorsque les valeurs extrêmes sont douteuses ou excessivement éloignées des autres valeurs de la série, on peut être amené à les retirer de l'étude et à calculer la moyenne sans en tenir compte. On parle alors de *moyenne élaguée*.

Il n'y a pas de règle pour choisir la moyenne élaguée plutôt que la moyenne. Dans notre exemple la moyenne élaguée serait la moyenne de la série $\{6,7,8\}$, soit 7, plus représentatif de la série.

La moyenne intéressera plutôt l'économiste.

1.4.2 Mesures de dispersion

Plusieurs caractères statistiques visent à indiquer comment sont dispersées les données de la série autour des valeurs centrales que sont le mode, la médiane et la moyenne, fournissant ainsi plus d'informations sur la série.

En Seconde, nous nous limiterons à ce qui suit.

Valeurs extrêmes : X_{min}, X_{max}

C'est la valeur minimale (X_{min}) et la valeur maximale (X_{max}) du caractère observé.

Ainsi, toutes les valeurs de la série appartiennent à l'intervalle $[X_{min} ; X_{max}]$.

Étendue : e

C'est la différence entre les valeurs extrêmes : $e = X_{max} - X_{min}$.

La comparaison des étendues de deux séries permet d'indiquer, sommairement, celle qui est la plus dispersée.

1.4.3 Résumé statistique

Le résumé statistique d'une série sera composé, en général, d'une ou plusieurs mesures centrales (en général moyenne ou médiane) et d'une mesure de dispersion (en général étendue).

L'objectif de la Seconde est de déterminer quand la médiane est plus pertinente que la moyenne.

On pourra adjoindre au résumé un graphique permettant de visualiser la série, parfois un diagramme en bâtons, le plus souvent un histogramme.

2 Activités

ACTIVITÉ 1

Vous trouverez en annexe page 8 trois extraits de relevés trimestriels de moyennes d'élèves de trois classes différentes.

Regroupez vous par 4 et, pour la matière que vous indiquera le professeur :

1. Arrondir les notes au demi point supérieur si besoin est. Ensuite, dans l'ordre qui vous paraît le plus judicieux :
   - calculer la moyenne ;
   - déterminer le(s) mode(s) ;
   - calculer l'étendue ;
   - déterminer la médiane ;
   - faire le diagramme en bâtons de la série ;
   - faire trois histogrammes à pas égaux : pas égal à 2, égal à 4 et égal à 10

   Les graphiques se feront avec en abscisses les notes de 0 à 20 (unité 1 cm) et en ordonnées les effectifs (unité 0,5 cm)
2. Afficher au tableau : moyenne, étendue, médiane, mode(s) de la matière étudiée. Comparer, commenter : peut-on résumer les résultats de la matière par la moyenne, la médiane, etc. ?
3. Peut-on retrouver la moyenne, l'étendue, le mode et la médiane pour l'ensemble des trois classes à partir de ce qui est affiché ? Si oui, comment ? Si non, que nous manque-t-il ?

ACTIVITÉ 2

Vous trouverez dans le tableau de la présente page le relevé journalier des précipitations (rosée, brouillard et pluie) à St Pierre de Chartreuse (montagne) et à St Étienne de St Geoirs (plaine), en mm. (*Données Météo France Saint-Martin d'Hérès*).

1. Pour chaque mois, calculer la moyenne et l'étendue des précipitations, puis la médiane.
   Quel résumé (moyenne étendue ou médiane étendue) semble le plus pertinent ? Justifier.
2. Calculer une moyenne élaguée des précipitations à St Étienne de St Geoirs (*une moyenne élaguée d'une série est la moyenne obtenue en privant cette série de certaines données que l'on considère comme aberrantes ; il n'y a pas de règle : ici, on enlèvera la donnée n°25*).
3. Comparer cette moyenne élaguée à la moyenne de St Pierre de Chartreuse de septembre 99.
   On dit que le massif de la Chartreuse est en général plus humide que la plaine. Qu'en pensez-vous ?

	St Pierre de Chartreuse			St Ét.
Date	sept 97	sept 98	sept 99	sept 99
1	0,8	0	0	0
2	21,2	17	0	0
3	0,2	0,5	0	3,4
4	0	59,4	0	0,2
5	0,2	2,6	0	0
6	0	1	10	3,8
7	0	15,4	0	0
8	0,2	1,2	0	0,2
9	0,2	0	0	0
10	0,2	34,8	0	0
11	0	45	0	0
12	31,6	55,6	0	0
13	19	44	0	0
14	0	9,6	11	0,2
15	0,2	14,4	1,6	0,8
16	0,2	0,6	0	0
17	0,2	0,2	1,6	0,8
18	0,2	0	0	0
19	0	0	42,4	27,2
20	0,2	0	0	0
21	0,2	0	9,8	0
22	1,2	5	0	0
23	0,2	0,2	1,4	0,4
24	0,2	0	1,2	0,2
25	0,2	0,6	51,8	189,2
26	0,2	15	1,2	0,8
27	0,2	5	14,2	11,2
28	0	10	6,8	1
29	0,2	4	0	0
30	0,2	28,4	44,4	23,6

ACTIVITÉ 3

Lors d'une enquête portant sur 1300 personnes, on a demandé le temps passé par jour devant le téléviseur. Les données relevées ont été regroupées par classe car les 1300 données sont en trop grand nombre pour être manipulées toutes ensembles. On a obtenu le tableau suivant :

Temps (h)	[0; 1[	[1; 2[	[2; 3[	[3; 4[	[4; 5[	[5; 6[	[6; 7]
Effectif	170	309	432	221	103	41	24

Les 1300 données ne sont alors plus accessibles dans le détail. Peut-on malgré tout obtenir de ce tableau les paramètres statistiques de la série (valeurs extrêmes, moyenne, médiane, etc.) ou, au moins, des valeurs approchées ?

- **Partie A : Valeurs extrêmes**
  1. Peut-on obtenir les valeurs minimale et maximale de la série ? Si oui les donner. Sinon donner une valeur (la plus grande possible) dont on est sûr qu'elle est plus petite que toutes les données de la série et une valeur (la plus petite possible) dont on est sûr qu'elle est plus grande que toutes les données de la série.
  2. Peut-on obtenir l'étendue de la série ? Si oui la donner. Sinon donner l'étendue maximale que peut avoir la série.

- **Partie B : Moyenne**
  1. Expliquer pourquoi on ne peut pas calculer la moyenne exacte du temps passé par jour devant la télévision.
  2. Pour obtenir une valeur approchée de la moyenne, on considère que toutes les données d'une classe sont égales au centre de la classe.
     Ainsi, par exemple, nous considèrerons que les 170 données de la classe $[0; 1[$ sont égales à $\frac{0+1}{2} = 0,5$.
     Compléter alors le tableau suivant et en déduire une valeur approchée de la moyenne :

Temps (h)	[0; 1[	[1; 2[	[2; 3[	[3; 4[	[4; 5[	[5; 6[	[6; 7]
Centre de la classe							
Effectif	170	309	432	221	103	41	24

- **Partie C : Médiane**
  1. Quel est le rang de la médiane de cette série ?
  2. Expliquer pourquoi on ne peut pas savoir le temps médian exact (la médiane exacte de cette série) passé par jour devant la télévision.
  3. Compléter le tableau suivant :

Temps passé devant la télévision inférieur à	1h	2h	3h	4h	5h	6h	7h
Effectif							

     Les effectifs obtenus dans la seconde ligne du tableau sont appelés effectifs cumulés croissants.
     Reporter les dans le tableau ci-dessous.

Temps passé devant la télévision inférieur à	1h	2h	3h	4h	5h	6h	7h
Effectif	170	309	432	221	103	41	24
Effectifs cumulés croissants							

  4. À l'aide de ces effectifs cumulés croissants et du 1., déterminer à quelle classe appartient la médiane exacte de cette série.
  5. Cette donnée est souvent trop approximative pour être utile en statistique et l'on a souvent besoin d'une estimation plus précise. On l'obtient avec un graphique.
     (a) Représenter le tableau sur un graphique en indiquant en abscisse les temps passés devant la télévision (1 h = 2 cm) et en ordonnée les effectifs cumulés croissants (100 = 1 cm).
     (b) À l'aide de ce graphique et du 1., déterminer une valeur approchée de la médiane de cette série.

- **Partie D : Mode**
  Lorsque les données sont regroupées par classes de même taille le mode n'est pas accessible mais la classe dont l'effectif et le plus grand est appelée classe modale. Lorsque les classes ne sont pas de la même taille, il existe des moyens d'estimer celle qui est modale, mais cette compétence n'est pas au programme de la Seconde.

Activité 4

1. Le tableau 1 présente la répartition des salaires mensuels dans une entreprise (*source : DoC TICE-MEN*).
Calculer la moyenne des salaires de cette entreprise et déterminer la médiane.
2. Une erreur a été commise dans les relevés : les effectifs correspondant aux salaires de 1100 € et 1400 € ont été permutés. Les données exactes sont celles du tableau 2.
Comparer les moyenne et médiane de cette série à celles de la précédente. Les variations étaient-elles prévisibles ?
3. Rêvons … 35 personnes ont un salaire de 3400 € et l'effectif total est inchangé. Utiliser le tableau 3 pour imaginer une répartition des effectifs telle que la médiane ne soit pas modifiée : comment va varier la moyenne ?
Quelle répartition des effectifs respectant ces contraintes donnera la moyenne des salaires la plus élevée (calculer alors cette moyenne) ? Quelle répartition des effectifs respectant ces contraintes donnera la moyenne des salaires la moins élevée (calculer alors cette moyenne) ?
Mathias prétend qu'avec 35 personnes ayant un salaire de 3400 €, la moyenne est obligatoirement plus élevée. A-t-il raison ?

Tableau 1

Salaire en €	Effectif	Effectif cumulé croissant
1 100	18	18
1 200	15	33
1 300	20	53
1 400	10	63
1 500	25	88
1 600	12	100
1 700	4	104
1 800	5	109
1 900	3	112
2 000	2	114
2 100	6	120
2 200	7	127
2 300	0	127
2 400	2	129
2 500	0	129
2 600	3	123
2 700	0	132
2 800	3	135
2 900	0	135
3 000	0	135
3 100	3	138
3 200	0	138
3 300	5	143
3 400	8	151

Tableau 2

Salaire en €	Effectif	Effectif cumulé croissant
1 100	**10**	
1 200	15	
1 300	20	
1 400	**18**	
1 500	25	
1 600	12	
1 700	4	
1 800	5	
1 900	3	
2 000	2	
2 100	6	
2 200	7	
2 300	0	
2 400	2	
2 500	0	
2 600	3	
2 700	0	
2 800	3	
2 900	0	
3 000	0	
3 100	3	
3 200	0	
3 300	5	
3 400	8	

Tableau 3

Salaire en €	Effectif	Effectif cumulé croissant
1 100		
1 200		
1 300		
1 400		
1 500		
1 600		
1 700		
1 800		
1 900		
2 000		
2 100		
2 200		
2 300		
2 400		
2 500		
2 600		
2 700		
2 800		
2 900		
3 000		
3 100		
3 200		
3 300		
3 400	**35**	**151**

Activité 5

1. Un professeur a noté des devoirs sur 40 ; la moyenne est $\overline{x} = 16$
   (a) Sur les copies, ce professeur a écrit les notes sur 20 ; quelle est la moyenne m des notes sur 20 ?
   (b) Trouvant la moyenne trop faible, le professeur décide de rajouter 1 point sur 20 à tous les élèves ; quelle est alors la moyenne X ?
   (c) Écrire X en fonction de x.
2. Calculer la moyenne de :
   - $x_1 = 10,0000432567189$
   - $x_2 = 10,0000432567370$
   - $x_3 = 10,0000432567127$
   - $x_4 = 10,0000432567433$
   - $x_5 = 10,0000432567156$
   - $x_6 = 10,0000432567238$
3. Pour calculer de tête la moyenne des notes 6 ; 7 ; 9 ; 12, des professeurs de mathématiques utilisent le procédé suivant :
   (a) On retranche 10 de chaque note ; on obtient respectivement -4 ; -3 ; -1 ; 2
   (b) La moyenne de ces nombres se calcule de tête facilement et est égale à $-1,5$
   (c) On ajoute 10 à ce résultat : on obtient 8,5 qui est la moyenne recherchée.

   Justifier cette méthode

3 Exercices

EXERCICE 1

Dans chaque cas, calculer la moyenne, le mode et la médiane de la série, vérifier que le narrateur dit la vérité et étudier quelles autres stratégies, s'il y en a, il aurait pu utiliser pour minimiser son résultat :

1. « Je n'ai eu que 8 sur 20 au contrôle de statistiques. Mes parents ne seront pas contents. Nous sommes 10 en classe. La meilleure note est 19. Ensuite il y a un 10, quatre 9, un 8 (moi) et trois 2. Je dirai donc que je suis au-dessus de la moyenne. »
2. « Encore un 8 ! Cette fois les notes sont 2, 3, 4, 5, 7, 8 (moi), 9, 9, 18, 19. Comment annoncer ma note ? Euh !... Je dirai que je suis dans la bonne moitié. »
3. « Toujours un 8 ! Cette fois il y a eu trois 7 et un 19, 18, 12, 11, 10, 8(moi) et 2. Tant pis, je dirai que je suis meilleur que le mode. »

EXERCICE 2

Voici trois séries de notes obtenues en mathématiques dans des classes de seconde (à effectifs très réduits) lors d'un contrôle sur les statistiques :

1. Déterminer la note médiane de chaque classe.
2. **SANS LA CALCULER**, conjecturer pour chaque série si la moyenne sera supérieure, inférieure ou proche de la médiane.

Classe 1	Classe 2	Classe 3
2	2	8
5	3	8
6	4	9
7	4	9
9	4	10
10	9	10
10	10	10
12	12	12
13	12	13
14	12	15
15	12	15
15	12	16
16	13	17
19	13	18

EXERCICE 3

Trois séries statistiques, comportant 10 données chacune, ont les paramètres suivants :

- Série A : Minimum 10 ; maximum 50 ; moyenne 28 ; médiane 20.
- Série B : Minimum 10 ; maximum 50 ; moyenne 30 ; médiane 30.
- Série C : Minimum 10 ; maximum 50 ; moyenne 21,5 ; médiane 25.

Conjecturer pour chacune de ces séries comment peuvent être réparties les données.

EXERCICE 4

Le recensement de 1999 a permis d'observer l'existence en France de dix communes de plus de 200 000 habitants. Le tableau ci-dessous donne la liste de ces communes et de leurs populations en milliers d'habitants.

1. Quelle est l'étendue de cette série statistique ? Que devient l'étendue quand on retire Paris de la liste ?
2. Comparer moyenne et médiane des populations de ces dix villes. Que constate-ton ? Comment expliquer ceci ?
3. Faire de même si on retire Paris de la liste. Que constatez-vous maintenant ?
4. Que choisiriez-vous entre « moyenne étendue » et « médiane étendue » pour résumer cette série statistique ? Expliquez votre choix.

Paris	2 116
Marseille	798
Lyon	445
Toulouse	391
Nice	341
Nantes	269
Strasbourg	264
Montpellier	225
Bordeaux	215
Rennes	206

EXERCICE 5

Dans une petite ville fictive où la taxe d'habitation est proportionnelle à la superficie de l'habitation, la répartition des habitations selon leur superficie est la suivante :

Superficie en m^2	Effectif
[10 ; 40[	14
[40 ; 70[	24
[70 ; 100[	54
[100 ; 120[	64
[120 ; 140[	32
[140 ; 170[	12

1. Déterminer une valeur approchée de la superficie moyenne des habitations de cette ville.
2. Un membre du conseil municipal propose d'exonérer la moitié des personnes : celles dont les habitations ont les superficies les plus faibles. Une personne dont l'appartement mesure 80 m^2 serait-elle exonérée ? Une personne dont l'appartement mesure 110 m^2 serait-elle exonérée ?
3. Un autre membre du conseil municipal propose d'exonérer le quart des personnes : celles dont les habitations ont les superficies les plus faibles. Une personne dont l'appartement mesure 80 m^2 serait-elle exonérée ?

EXERCICE 6

Dans deux entreprises A et B, les employés sont classés en deux catégories : ouvriers et cadres.

Les deux tableaux qui suivent donnent la répartition des employés en fonction de leur catégorie professionnelle et de leur salaire mensuel net, en euros. On suppose qu'à l'intérieur de chaque classe, la répartition est régulière.

1. (a) Calculer la moyenne des salaires de tous les employés de l'entreprise A
   (b) Calculer la moyenne des salaires des ouvriers de l'entreprise A
   (c) Calculer la moyenne des salaires des cadres de l'entreprise A
2. Faire les mêmes calculs pour l'entreprise B
3. Le P.D.G. de l'entreprise A dit à celui de l'entreprise B : « Mes employés sont mieux payés que les vôtres. »
   « Faux » répond ce dernier, « mes ouvriers sont mieux payés et mes cadres également. »
   Expliquer ce paradoxe.

Entreprise A

Salaires	$[1\,000 ; 2\,000[$	$[2\,000 ; 3\,000[$	$[3\,000 ; 4\,000[$
Ouvriers	114	66	0
Cadres	0	8	12

Entreprise B

Salaires	$[1\,000 ; 2\,000[$	$[2\,000 ; 3\,000[$	$[3\,000 ; 4\,000[$
Ouvriers	84	42	0
Cadres	0	12	12

EXERCICE 7

Lors d'une étude d'une population de rats, K. Miescher a observé l'évolution d'une population de 144 rats. Le tableau suivant indique la durée de vie (en mois) des rats.

Ainsi, un seul rat a vécu entre 10 et 15 mois, trois ont vécu entre 15 et 20 mois, neuf entre 20 et 25 mois etc. On suppose que, dans chaque classe, la répartition est régulière.

1. Évaluez l'étendue de cette série
2. Évaluez la moyenne de la durée de vie d'un rat dans cette population
3. Quelle est la classe médiane de la durée de vie d'un rat dans cette population. On prendra le centre de cette classe comme valeur approchée de la médiane. Combien vaut alors cette valeur approchée de la médiane ?
4. En observant la moyenne et la médiane, quel commentaire peut-on faire ?

Durée de vie (en mois)	Effectif
$[10 ; 15[$	1
$[15 ; 20[$	3
$[20 ; 25[$	9
$[25 ; 28[$	12
$[28 ; 30[$	13
$[30 ; 32[$	20
$[32 ; 34[$	23
$[34 ; 36[$	26
$[36 ; 38[$	22
$[38 ; 40[$	11
$[40 ; 42[$	3
$[42 ; 43[$	1

Annexes

TAB. 1 – Relevés de notes pour l'activité 1

Classe 1 (Terminale ES)

Philo	H.-G.	S.E.S.	Maths	E.P.S.
6,6	11,5	9,5	12,3	10
7,6	10,3	9,7	5,3	14,4
9,7	12,7	10,3	12,2	8
8,1	10,4	11	12,1	6
6,5	10,7	10,2	8,9	10
9,4	12,3	10	10	17,2
7,5	10,2	13,2	7,4	11,5
10,8	11,5	11	7,2	13
8,6	9,7	9,5	11,2	10
9	10,8	9,5	12,8	12,9
7,2	10	11,3	12,6	8
10,3	7,3	10,3	11,4	10
9,7	9,8	9	9,7	13,3
8	12	9	10	9
10,3	13,9	10,5	11	6
8,6	10,1	8,8	9,6	15
6,5	11	6,8	7,1	7,4
13,3	13,3	15,5	16,7	15
7,4	8,8	12,3	12,8	15,4
6,7	NN	10,8	5,3	7,7
9,1	12,9	10,5	7,5	Disp.
6,9	7,8	10,3	9,9	11
8,4	11,6	10,8	6	14,4
10,3	14,7	9	8,4	6
9,8	7,9	11,5	11,9	16
10,6	10,7	12,7	6,7	13,7
9	10,5	11,5	10,9	12
11,4	10,6	13,2	9,6	10
6,8	8,8	10,2	7,8	Disp.

Classe 2 (Première ES)

Fran	H.-G.	S.E.S.	Maths	E.P.S.
8,25	8,75	13	10	14,5
13	12	13,5	13,5	19,5
12	9,5	12,5	7	17,5
10,75	8,5	13	8,5	16
11,75	11	12	12	15
10,25	9,75	13	11,5	15
10,25	7	10	8	16,5
9,75	8,75	12	8,5	17,5
11	8,25	11,5	7,5	17
12	10	9	12,5	14
7,25	9	12	10	17
6,5	7,25	11	8	14
10	8,75	9	4,5	18
9,25	8,25	13	10,5	8
11,75	6,75	12,5	11,5	NN
8,75	9,5	12	7,5	16
11	9	12	5	NN
12,25	9	12	5	NN
9,25	10,75	10	8	18
12	9,25	11,5	6	15,5
8,25	7,75	11	9,5	17
12,5	10,25	13	13,5	10
9,75	12,25	9	6,5	17
11,75	9,25	10,5	8	NN
11,25	6,5	15	10,5	14
12,75	10,75	10	10	6,5
12,25	11	10	8	14,5
12,25	11,25	10	10	16
10	10,75	11	9,5	14,5
11	10,25	NN	9	NN

Classe 3 (Seconde)

Fran	H.-G.	P.-C.	S.V.T.	E.P.S.
10,5	11,25	11	8,5	17
9	9	6	8,5	14
9,5	14,5	11,5	8,25	14
12,25	10,5	10,5	11	11
11,25	11,25	13,5	12	14
8	10	11	9,75	13
8,75	9	10	6,25	15
7	7	7	8	11
8,5	15	8	7,25	14
7,5	10	11	9,5	15
12,2	13,5	12,5	13	12
11,25	10,75	13	13	12
16,5	19	18	14,75	11
15	15	16	15,5	12
13,25	18	18,5	15,5	14
12,5	16	15,5	13	12
8,5	11,25	12,5	1,25	13
10	13,5	14,5	13,5	NN
8,5	10,5	7,5	11,25	13
12	14,5	16	10,5	14
10,75	11,25	9,5	10	12
10	11	12,5	13	10
10,25	6,5	9,5	9,75	13
7,75	9	9,5	6,25	14
12	13	12	13	14
10,25	12,25	12	13,5	11