

Séquence IV – Statistiques

Exercice IV.1

Voici la répartition des notes à un test :

Notes	4	6	8	9	10	11	12	14	16
Effectifs	13	23	28	10	13	11	13	8	1

Calculer la moyenne des notes.

Exercice IV.2

Un chef de rayon commande des charentaises chez deux grossistes : 60 % de la commande chez l'un, au prix de 25 € la paire, et 40 % restant chez l'autre, au prix de 17 € la paire. Calculer le prix moyen des charentaises.

Exercice IV.3

En 2015, aux épreuves anticipées du Bac, en première, il y avait deux épreuves anticipées, en français et en histoire-géographie. Le coefficient de la note de français était 2 et celui d'histoire-géographie 3.

Calculer la note moyenne, pour chacune des personnes suivantes :

1. Pour Rachel qui a obtenu 12 en français et 7 en histoire-géographie.
2. Pour Maxime qui a obtenu 8 en français et 13 en histoire-géographie.
3. Pour Tony qui a obtenu 6,5 en français et 14,5 en histoire-géographie.

Exercice IV.4

On a ordonné les valeurs de chacune des séries suivantes. Calculer la moyenne et donner la médiane dans chacun des cas suivants :

1. Les dix tailles de chaussures d'une famille :

37 – 37 – 39 – 39 – 39 – 41 – 42 – 42 – 45 – 45

2. Le nombre d'enfants parmi les douze familles d'un immeuble :

0 – 0 – 0 – 0 – 1 – 1 – 2 – 2 – 3 – 3 – 3 – 4

3. Le salaire des sept salariés d'une PME (en euros) :

930 – 1 023 – 1 147 – 1 250 – 1 403 – 1 810 – 2 365

Exercice IV.5

Dans un lycée, on étudie les moyennes du premier trimestre de deux classes, appelées respectivement Jaune et Rouge.

Partie A :

Les 25 élèves de la classe Jaune ont obtenu les moyennes suivantes :

3 ; 4 ; 5 ; 7 ; 7 ; 10 ; 10 ; 10 ; 10 ; 10 ; 11 ; 11 ; 12 ; 12 ; 12 ; 12 ; 13 ; 13 ; 13 ; 14 ; 15 ; 15 ; 16 ; 18

1. Déterminer la médiane M_e , le premier quartile Q_1 et le troisième quartile Q_3 de cette série statistique.
2. Représenter le diagramme en boîte correspondant en faisant apparaître les valeurs extrêmes.
3. Calculer la moyenne de la classe Jaune.

Partie B :

Les indicateurs de la classe Rouge permettant de résumer la série statistique des moyennes du premier trimestre sont les suivants :

- Minimum = 3
- $Q'_3 = 12$
- $Q'_1 = 8$
- Maximum = 17
- $M'_e = 10$

1. Représenter le diagramme en boîte correspondant en dessous de celui de la classe Jaune.
2. Parmi les affirmations suivantes, lesquelles sont vraies, fausses ou indécidables ? (indécidable signifie que l'on ne peut pas conclure avec les éléments en présence). Justifier la réponse dans chacun des cas.

- (a) 50 % des élèves de la classe Rouge ont une note comprise entre 10 et 12.
- (b) 75 % des élèves de la classe Rouge ont une note inférieure ou égale à 12.
- (c) Au moins 50 % des élèves de la classe Rouge ont une note inférieure ou égale à la note médiane de la série de la classe Jaune.

Exercice IV.6

On relève la température d'une forêt chaque heure pendant quatre jours.

Les 97 résultats obtenus ont été triés et sont rassemblés dans le tableau suivant :

Température (en °C)	14,5	15	15,5	16	16,5	17	17,5	18	18,5	19	19,5
Nombre de fois où cette température a été relevée	5	7	10	12	15	10	11	9	7	7	4

1. (a) Déterminer la médiane M_e et les quartiles Q_1 et Q_3 de cette série statistique.

On appelle **premier décile** (noté D_1) la plus petite valeur de la température telle qu'au moins 10 % des valeurs sont inférieures ou égales à D_1 . On appelle **neuvième décile** (noté D_9) la plus petite valeur telle qu'au moins 90 % des valeurs lui sont inférieures ou égales.

(b) Justifier que $D_1 = 15$ et calculer D_9 .

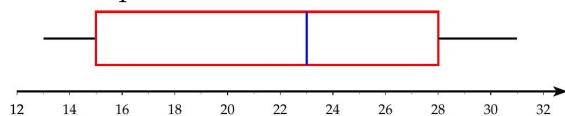
(c) Calculer l'écart interquartile.

2. La température a été relevée de la même manière et aux mêmes instants dans un champ à l'extérieur de la forêt. Cette deuxième série de résultats ne figure pas ici, mais :

- la médiane de cette deuxième série est $M' = 23^\circ\text{C}$
- les quartiles de cette deuxième série sont $Q'_1 = 15^\circ\text{C}$ et $Q'_3 = 28^\circ\text{C}$
- les déciles de cette deuxième série sont $D'_1 = 13^\circ\text{C}$ et $D'_9 = 31^\circ\text{C}$.

(a) Calculer l'écart interquartile de cette nouvelle série.

(b) On considère un diagramme en boîte de cette série. Les extrémités du diagramme correspondent aux premier et neuvième déciles. Construire celui de la série des températures relevées dans la forêt.



(c) En quelques lignes, expliquer quelle semble être l'influence des arbres sur la température à l'intérieur de la forêt.

Exercice IV.7

On connaît la répartition des notes à un test noté sur 10 :

Notes	2	3	4	5	6	7	8
Effectifs	13	41	28	13	13	8	1

Calculer la moyenne et l'écart-type correspondant à cette série.

Exercice IV.8

Une machine est réglée pour produire des paquets de pâtes de 500 grammes. Pour vérifier ce réglage, on prélève un lot de 100 paquets de pâtes que l'on pèse. Les résultats sont donnés dans le tableau ci-dessous :

Masse (en grammes)	Nombre de paquets
495	1
496	4
497	10
498	12
499	20
500	20
501	16
502	9
503	5
504	2
505	1

- Déterminer la masse moyenne d'un paquet $\bar{x}$ ainsi que l'écart-type σ de la série.
- D'après le protocole de maintenance de la machine, il faut procéder à un réglage si l'intervalle $[\bar{x} - 2\sigma ; \bar{x} + 2\sigma]$ contient moins de 95 % des paquets produits. D'après les mesures, faut-il régler la machine ?

Exercice IV.9

1. L'affirmation suivante est-elle vraie ?

« Si deux séries ont exactement la même moyenne et le même écart-type, alors ces séries ont exactement les mêmes valeurs »

2. Énoncer la réciproque de l'affirmation précédente et indiquer si cette réciproque est vraie.

Exercice IV.10

Indiquer si les affirmations suivantes sont vraies et justifier.

- Si l'on divise toutes les valeurs d'une série par 2, alors la variance est divisée par 2.
- Plus une série a de valeurs, plus l'écart-type est grand.



Exercice IV.11

Partie 1 - Conjecture

On considère l'enquête statistique suivante :

« Nous aimerions recueillir vos impressions concernant nos services. Veuillez indiquer votre note du niveau de satisfaction en utilisant l'échelle suivante :

- 1 - Pas du tout satisfait
- 2 - Moyennement satisfait
- 3 - Satisfait
- 4 - Très satisfait
- 5 - Extrêmement satisfait.

Cinq personnes ont répondu, les résultats de l'enquête sont les suivants : 2 personnes sont satisfaites, 1 personne est très satisfaite et 2 personnes sont extrêmement satisfaites. »

1. Déterminer la moyenne des notes attribuées à l'entreprise.
2. On considère la fonction f définie, pour tout $x \in \mathbb{R}$, par :

$$f(x) = \frac{1}{5} \left(2(x-3)^2 + 1(x-4)^2 + 2(x-5)^2 \right)$$

Déterminer pour quelle valeur de x la fonction f admet un minimum sur $\mathbb{R}$. Que vaut alors ce minimum ?

3. Quelle conjecture peut-on émettre ?

Partie 2 - Démonstration

On considère la série statistique suivante :

- Les réponses données par les personnes interrogées (les valeurs du caractère) sont les réels $x_1, x_2, \dots, x_k$, avec k un entier supérieur à 1.
- Les effectifs correspondants sont notés $n_1, n_2, \dots, n_k$. Ainsi, avec ces notations, la réponse x_1 est donnée n_1 fois, etc.

On note $N = \sum_{i=1}^k n_i$. Soit f la fonction définie, pour tout $x \in \mathbb{R}$, par

$$f(x) = \frac{1}{N} \left(n_1(x-x_1)^2 + \dots + n_k(x-x_k)^2 \right)$$

1. Soit $x \in \mathbb{R}$. Exprimer $f(x)$ en utilisant le symbole $\sum$
2. Montrer que f admet un minimum sur $\mathbb{R}$, que ce minimum est atteint en la moyenne pondérée des valeurs du caractère et qu'il est égal à la variance de la série statistique.



Exercice IV.12

Le « jeu de la boule » est constitué d'un plateau cylindrique fixe comprenant une couronne de neuf numéros allant 1 à 9.

En théorie, chacun des neuf chiffres a autant de chance d'être obtenu à chaque tirage.

Un casino fait l'objet d'un contrôle suite à des plaintes de joueurs. En effet, ils estiment que la boule ne tourne pas rond ! Ainsi, la répression des fraudes a effectué un grand nombre de lancers et a obtenu les résultats suivants :

Numéros	1	2	3	4	5	6	7	8	9
Effectifs	98	97	101	96	131	132	125	107	113

1. Calculer les fréquences d'apparition de chacun des neuf chiffres.
2. Pour mesurer l'écart entre la fréquence théorique de $\frac{1}{9}$ et les fréquences obtenues, on introduit le coefficient

$$d_{\text{obs}}^2 = \sum_{i=1}^9 \left(f_i - \frac{1}{9} \right)^2$$

Calculer $9000d_{\text{obs}}^2$ (on a multiplié par 9000 pour avoir un résultat moins petit).

3. Idéalement, $d_{\text{obs}}^2 = 0$. Mais dans la pratique, il y a quelques écarts entre les fréquences théoriques de $\frac{1}{9}$ et les fréquences obtenues. Donc $d_{\text{obs}}^2 > 0$, mais d_{obs}^2 ne doit pas être trop grand...

Pour avoir une idée des valeurs de d_{obs}^2 qu'on peut obtenir avec un plateau parfaitement équilibré, on a effectué 1000 simulations de 1000 lancers avec un ordinateur et calculé, à chaque fois, la valeur de d^2 obtenue.

Voici la répartition des résultats :

$9000d^2$	$[0; 2[$	$[2; 4[$	$[4; 6[$	$[6; 8[$	$[8; 10[$	$[10; 12[$
Effectifs	17	106	216	208	175	112

$9000d^2$	$[12; 15[$	$[15; 20[$	$[20; 25[$	$[25; 30[$	$[30; 60[$
Effectifs	100	50	12	3	1

À partir de cette série statistique :

- (a) Calculer la valeur médiane ;
 - (b) Calculer les premier et troisième quartiles ;
 - (c) Calculer les premier et neuvième déciles ;
 - (d) Tracer le diagramme en boîte.
4. On considère que le jeu n'est pas truqué si la valeur de $9000d_{\text{obs}}^2$ calculée à la question 2 est inférieure au neuvième décile de la série statistique de la question 3. Selon ce critère, le jeu est-il truqué ?