

Corrigé Statistiques

Les paramètres pour comparer des séries statistiques

Exercice 1.

Le tableau ci-dessous donne le nombre de livres lus lors des 12 derniers mois par les élèves de trois classes. Comparer les en utilisant les paramètres statistiques.

Nombre	0	1	2	3	4	5	6	7	8
Classe 1	0	1	4	8	0	1	5	2	0
Classe 2	3	5	4	3	2	7	1	5	3
Classe 3	0	0	1	1	0	1	1	5	5

Classe 1

Effectif total : $n = 1 + 4 + 8 + 1 + 5 + 2 = 21$

$$\text{Moyenne : } \bar{x} = \frac{1 \times 1 + 2 \times 4 + 3 \times 8 + 5 \times 1 + 6 \times 5 + 7 \times 2}{21} \approx 3,905$$

Minimum : 1. C'est la plus petite valeur.

Quartile 1 : 3. $\frac{21}{4} = 5,25$, c'est donc la 6ème valeur (on arrondit à l'entier supérieur). La calculatrice utilise la définition internationale et trouve 2,5, c'est normal.

Médiane : 3. $\frac{21}{2} = 10,5$, c'est donc la 11ème valeur (on ajoute 0,5).

Quartile 3 : 6. $\frac{3 \times 21}{4} = 15,75$, c'est donc la 16ème valeur (on arrondit à l'entier supérieur).

Maximum : 7. C'est la plus grande valeur.

Mode : 3. C'est la valeur la plus courante.

Écart type : $\sigma \approx 1,849$. Obtenu à la calculatrice.

Étendue : $7 - 1 = 6$. Différence entre maximum et minimum.

Écart inter-quartiles : $6 - 3 = 3$. Différence entre Q3 et Q1.

Classe 2

Effectif total : $n = 3 + 5 + 4 + 3 + 2 + 7 + 1 + 5 + 3 = 33$

Moyenne : $\bar{x} \approx 3,939$

Minimum : 0.

Quartile 1 : $2. \frac{33}{4} = 8,25$, c'est donc la 9ème valeur.

Médiane : $4. \frac{33}{2} = 16,5$, c'est donc la 17ème valeur.

Quartile 3 : $6. \frac{3 \times 33}{4} = 24,75$, c'est donc la 25ème valeur.

Maximum : 8.

Mode : 5.

Écart type : $\sigma \approx 2,558$.

Étendue : $8 - 0 = 8$.

Écart inter-quartiles : $6 - 2 = 4$.

Classe 3

Effectif total : $n = 1 + 1 + 1 + 1 + 5 + 5 = 14$

Moyenne : $\bar{x} \approx 6,5$

Minimum : 2.

Quartile 1 : $6. \frac{14}{4} = 3,5$, c'est donc la 4ème valeur.

Médiane : $7. \frac{14}{2} = 7$, il faut créer une valeur entre la 7ème et la 8ème.

Quartile 3 : $8. \frac{3 \times 33}{4} = 10,5$, c'est donc la 11ème valeur.

Maximum : 8.

Mode : 7 et 8.

Écart type : $\sigma \approx 1,842$.

Étendue : $8 - 2 = 6$.

Écart inter-quartiles : $8 - 6 = 2$.

Comparaison

La classe 2 a le plus grand effectif, le double environ de la classe 3 qui a le plus faible effectif.

En terme de niveau c'est la classe 3 qui est supérieure avec une moyenne bien plus grande que les deux autres mais également un mode et des paramètres de position (extremums, quartiles, médiane) supérieurs.

Entre la classe 1 et 2 c'est moins tranché mais c'est à l'avantage de la 2. La 1 a un meilleur minimum et quartile 1. La 2 a une meilleure moyenne, médiane, maximum et mode.

La 3 est légèrement plus homogène que la 1 elle même plus homogène que la 2. Les trois paramètres de dispersion (écart type, étendue et écart inter-quartiles) montrent cela.

Exercice 2.

Le tableau ci-dessous donne l'âge des membres de trois clubs de sport. Comparer les en utilisant les paramètres statistiques.

Nombre]10;18]]18;30]]30;40]]40;50]]50;60]	60 et +
Club 1	0	0	40	10	30	20
Club 2	15	5	10	8	15	5
Club 3	15	20	15	10	0	20

La première difficulté dans ce tableau consiste à avoir un âge précis pour chaque colonne. De manière pragmatique on peut considérer que la classe d'âge]10;18]

a globalement un âge de 14 (la moyenne). On procède de même pour les autres classes d'âge en mettant 24, 35, 45, 55 mais la dernière pose à nouveau problème. Je choisis ici 70 pour représenter la classe d'âge de 60 ans et plus mais on peut faire un autre choix.

Club 1

Effectif total : $n = 100$

Moyenne : $\bar{x} = 49$

Minimum : 35.

Quartile 1 : 35.

Médiane : 50.

Quartile 3 : 55.

Maximum : 70.

Mode : 35.

Écart type : $\sigma \approx 13,4$.

Étendue : 35.

Écart inter-quartiles : 20.

Club 2

Effectif total : $n = 58$

Moyenne : $\bar{x} \approx 38,2$

Minimum : 14.

Quartile 1 : 14.

Médiane : 35.

Quartile 3 : 55.

Maximum : 70.

Mode : 14 et 55.

Écart type : $\sigma \approx 18,4$.

Étendue : 56.

Écart inter-quartiles : 41.

Club 3

Effectif total : $n = 80$

Moyenne : $\bar{x} \approx 38,3$

Minimum : 14.

Quartile 1 : 24.

Médiane : 35.

Quartile 3 : 45.

Maximum : 70.

Mode : 24 et 70.

Écart type : $\sigma \approx 20,5$.

Étendue : 56.

Écart inter-quartiles : 21.

Comparaison

Les paramètres montrent clairement que le club 1 a des âges globalement plus éle-

vés et plus homogènes que les deux autres. Il y a également plus de membres.

Entre le club 2 et 3 c'est plus disputé. L'âge moyen, le quartile 1 et le mode sont un peu plus élevés dans le club 3, le quartile 3 est meilleur dans le club 2. Côté homogénéité c'est très serré : l'écart type est meilleur dans le club 2, l'écart inter-quartiles meilleur dans le club 3.

Exercice 3.

1. Dans l'exercice 1, si tous les élèves lisent un livre supplémentaire qu'arrive-t-il aux différents paramètres statistiques et quel est l'impact sur nos comparaisons?

Si tous les élèves lisent un livre de plus :

La moyenne, les extremums, les quartiles, la médiane et le mode augmentent de 1.

L'effectif total, l'écart type, l'étendue et l'écart inter-quartiles ne changent pas.

Cela ne change rien aux comparatifs.

2. Toujours dans l'exercice 1, si tous les élèves lisent le double de livres qu'arrive-t-il aux différents paramètres statistiques et quel est l'impact sur nos comparaisons?

Dans ce cas tous les paramètres doublent sauf l'effectif total qui ne change pas.

Cela ne change rien aux comparatifs.

3. Dans l'exercice 2 que deviendront les paramètres statistiques dans deux ans si les membres des différents clubs ne changent pas?

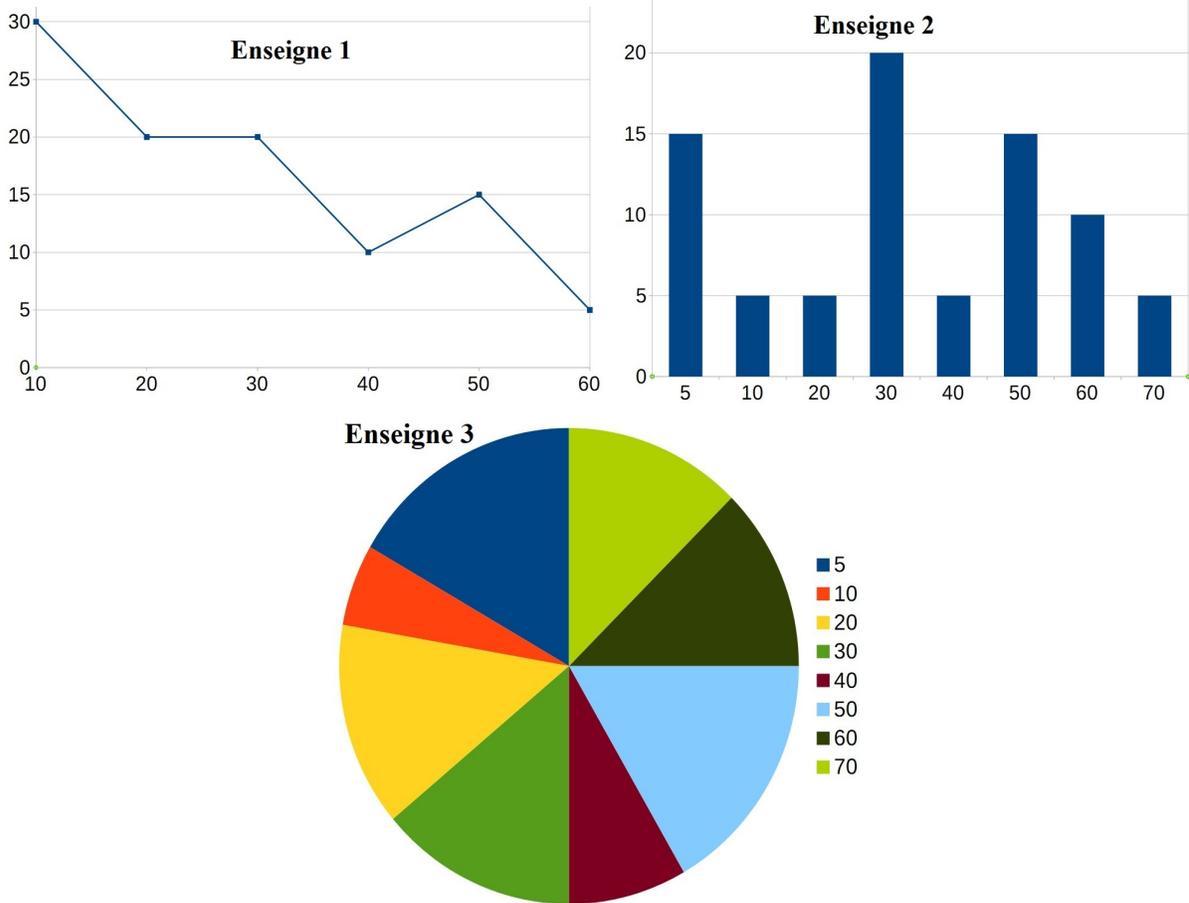
La moyenne, les extremums, les quartiles, la médiane et le mode augmentent de 2.

L'effectif total, l'écart type, l'étendue et l'écart inter-quartiles ne changent pas.

Cela ne change rien aux comparatifs.

Exercice 4.

Les graphiques ci-dessous donnent le prix de vente de jeux vidéo vendus dans trois enseignes. Comparer les en utilisant les paramètres statistiques.



Commençons par faire un tableau avec les effectifs de ventes de chaque enseigne.

Prix	5	10	20	30	40	50	60	70
Enseigne 1	0	30	20	20	10	15	5	0
Enseigne 2	15	5	5	20	5	15	10	5
Enseigne 3	60°	20°	50°	50°	30°	60°	45°	45°

Afin de simplifier la visibilité nous allons également donner les paramètres dans un tableau. L'enseigne 3 ne permet pas de savoir l'effectif total et l'enseigne 1 est probablement en % donc nous n'allons pas mettre ce paramètre.

	Enseigne 1	Enseigne 2	Enseigne 3
Moyenne	27,5	34,1	36,2
Minimum	10	5	5
Quartile 1	10	15	20
Médiane	25	30	35
Quartile 3	40	50	55
Maximum	60	70	70
Mode	10	30	5 et 50
Écart type	15,8	20,6	22,1
Étendue	50	65	65
Écart inter-quartiles	30	35	35

Comparaison

On constate que l'enseigne 3 a des tarifs globalement supérieurs aux deux autres (Moyenne, Q1, médiane, Q3 et maximum). Entre l'enseigne 1 et 2 c'est plutôt la 2 qui est plus chère (moyenne, Q1, médiane, Q3, maximum).

Côté homogénéité des tarifs la 1 est plus homogène (les paramètres de dispersion sont plus petits) puis la 2 puis la 3.