**Statistiques descriptives. Analyse de données.**

1. **Moyenne**
2. **Moyenne et moyenne pondérée**

**DÉFINITION**

Soit une série statistique de valeurs, d’effectifs respectifs donnés dans le tableau ci-dessous.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Valeur |  |  |  |  |
| Effectif |  |  |  |  |

La **moyenne pondérée** de la série statistique ci-contre est le nombre réel, noté , tel que :

**Exemple**

Audrey prend souvent le train venant de Rouen en direction de Paris. En rentrant dans le wagon, elle compte le nombre de places assises disponibles.

Après 20 trajets, elle obtient les résultats ci-contre.



Le nombre moyen de places assises disponibles sur ces 20 trajets est :

**Remarque**

Dans la propriété précédente, en posant , on obtient la fréquence des :

  ; celle de , etc.

On en déduit donc la formule

**Propriété - Moyenne pondérée**

On considère une série statistique constituée de valeurs affectées de coefficients (ou poids) .

La moyenne pondérée de cette série est

**Remarque**

La formule est la même que pour la moyenne d'une série donnée sous forme de tableau d'effectifs, mais il est important de comprendre que les séries sont de natures différentes :

* dans le 1er cas (avec un tableau d'effectifs) la série est constituée de valeurs, précisément :  ;  ; … ;
* dans le 2ème cas (avec les coefficients ou poids) la série est constituée de valeurs (éventuellement identiques) auxquelles on attribue un coefficient (ou poids) qui correspond à l'importance de la valeur.

**Exemple**

Ce trimestre, Émilia a eu quatre contrôles de mathématiques (notés sur 20) de coefficients 1 ; 1,5 ; 4 et 0,5 auxquels elle a obtenu respectivement les notes 8 ; 9 ; 20 et 5.

Sa moyenne en mathématiques est donc

**Remarque**

On constate que, bien qu'elle ait eu 3 notes sur 4 en dessous de 10, sa moyenne est bonne : ceci est dû au fait qu'elle ait eu une très bonne note (20) à un devoir ayant un grand « poids » (donc une plus grande importance) par rapport aux autres.

* **Exercice résolu 1 et 2 page 317**











1. **Linéarité de la moyenne**

**Propriété – Linéarité de la moyenne**

Soit et deux nombres réels et une série statistique de moyenne .

* Si on multiplie par toutes les valeurs de la série, on obtient la moyenne de la nouvelle série en ……………………………………………………………………………………………..

Autrement dit, la moyenne de la série est ………………

* Si on ajoute à toutes les valeurs de la série, on obtient la moyenne de la nouvelle série en ……………………………………………………………………………………………..

Autrement dit, la moyenne de la série est ………………

* Les deux points précédents assurent également que la moyenne de la série

 est………………………

**Exemple**

La semaine dernière, pour se préparer le matin, Juan a mis 20 minutes en moyenne :

* S’il avait mis deux minutes de plus chaque jour, il aurait mis en moyenne pour se préparer ;
* S’il avait mis 5 % de temps en plus chaque jour, c'est-à-dire si son temps de préparation avait été multiplié par chaque jour, alors il aurait mis en moyenne minutes pour se préparer.

**Remarque**

La propriété de linéarité de la moyenne reste vraie lorsque :

* on ………………………. un même nombre à toutes les valeurs de la série, puisque soustraire un nombre est équivalent à ……………………………… ;
* on ………………… toutes les valeurs de la série par un même nombre, puisque diviser par un nombre est équivalent à …………………………………..

**Exemples**

① Si l'on soustrait 5 à toutes les valeurs d'une série statistique de moyenne alors cela revient à ………... toutes ses valeurs. Ainsi, la moyenne de la nouvelle série est .

② Si l'on divise par 4 toutes les valeurs d'une série statistique de moyenne alors cela revient à …………………….. toutes ses valeurs. Ainsi, la moyenne de la nouvelle série est

* **Exercice résolu 3 page 319**





1. **Écart-type**

**Définition - Écart-type**

L'écart-type s d'une série statistique est un **……………………….** de cette série statistique autour de la moyenne. Concrètement il donne une certaine mesure de ……………………………………………………

……………………………………………………………… :

* plus l'écart-type s d'une série est petit, plus les valeurs de la série sont ……………………………….

…………………………………………….., donc plus la série est …………………………..;

* plus l'écart-type s d'une série est grand, plus les valeurs de la série sont ……………………………..., donc moins la série est ………………………………..

**Remarque**

Nous utiliserons la calculatrice pour déterminer l'écart-type (voir Tuto Vidéo et le TP1) mais il existe des formules pour le calculer (pour info : livre page 292).

**Exemple**

On considère deux entreprises de 10 employés :

* l'entreprise 1 dans laquelle 5 employés gagnent 2 500 € et 5 employés gagnent 3 500 € par mois ;
* l'entreprise 2 dans laquelle 9 employés gagnent 1 200 € et 1 employé gagne 19 200 € par mois.

Le salaire moyen dans l'entreprise 1 est de

 et le salaire moyen dans l'entreprise 2 est de .

On comprend pourtant bien que la répartition des salaires dans les deux entreprises est extrêmement différente : la moyenne seule ne fournit pas une information suffisante pour résumer la série de manière satisfaisante.

On utilise alors un indicateur permettant de mesurer l'homogénéité des salaires dans les deux entreprises : l'écart-type.

Avec la calculatrice, on obtient :

* l'écart-type de la série des salaires de l'entreprise 1 , qui est ………………………
* l'écart-type de la série des salaires de l'entreprise 2, qui est de ………………………..

On constate donc que, bien que le salaire moyen soit le même dans les deux entreprises, 3 000 € :

* dans l'entreprise 1, les employés ont globalement des salaires ……………………………………. ;
* dans l'entreprise 2, les employés ont globalement des salaires ……………………………………..

**Remarques**

* On peut utiliser le couple (moyenne ; écart-type) pour résumer une série et en comparer plusieurs.
* Les valeurs éloignées de la moyenne ont de l'influence sur l'écart-type, plus précisément elles le font augmenter.
* Pour les séries dont le diagramme en bâtons (ou l'histogramme) est en forme de cloche, on peut s'attendre à trouver l'essentiel des valeurs de la série (environ 95 %) dans l'intervalle

 où m désigne la moyenne et s l'écart-type de la série (voir le TP2)

* **Exercice résolu 4 page 319**





1. **Quartiles et écart interquartile**

**DÉFINITIONS - Quartiles**Les valeurs d'une série statistique étant **rangées par ordre croissant**

* **le premier quartile** est la plus petite valeur de la série telle qu'**au moins 25 %** des valeurs de la
série sont **…………………………………..** ,
* **le troisième quartile** est la plus petite valeur de la série telle qu'**au moins 75 %** des valeurs de la
série sont **…………………………………..**.

**Exemple**

On considère la série ordonnée de 9 valeurs . On a alors :

Plus de 25 % des valeurs : .

Moins de 25 % des valeurs : .

Plus de 75 % des valeurs :

Moins de 75 % des valeurs :

**Propriété - Rang des quartiles**

Pour une série ordonnée d'effectif , (resp. ) est la -ième valeur où est le plus petit entier supérieur ou égal à (resp. ).

**Exemple**

On reprend la série du nombre de places disponibles dans un train sur 20 trajets du paragraphe 1. a.



* Pour trouver , on calcule donc est la …… valeur, c’est-à-dire (car la série est
* Pour trouver , on calcule donc est la …….. valeur, c’est-à-dire .

**Propriété - Médiane**

Pour une série ordonnée d'effectif , la médiane est :

* la valeur de rang si est …………….. ;
* la moyenne des valeurs de rang et si est ……..

**Exemple**

Dans l'exemple précédent, la série a pour effectif qui est pair. et donc la médiane est la moyenne des ……. et …….. valeurs, c'est-à-dire

**Remarque**

Pour une série statistique de valeur minimale et de valeur maximale , chacun des intervalles , , et contient au moins 25 % des valeurs de la

série (et environ 25 % si la série est de grand effectif et constituée essentiellement de valeurs différentes).



**Définition - Écart interquartile**

L'écart interquartile d'une série statistique est . Il s'agit d'un indicateur de dispersion.

**Remarques**

* Plus l'écart interquartile est petit, plus les valeurs « centrales » de la série (celles dans l'intervalle ) sont proches les unes des autres. Les valeurs supérieures à ou inférieures à n'ont pas d'influence sur l'écart interquartile.
* On peut utiliser le couple (médiane ; écart interquartile) pour résumer une série et en comparer plusieurs.
* **Exercice résolu 5 page 321**





