Séquence 8 :

Statistique descriptive

I. Moyenne et écart-type

Activité 1 p. 302 : découverte de la méthode de calcul de la moyenne pondérée.

A. Moyenne pondérée :

Définition : On considère la série statistique $(x_1; n_1)$, $(x_2; n_2)$, … , $(x_p; n_p)$ où x_1, x_2, … x_p sont les valeurs du caractère étudié et n_1,n_2, … , n_p les effectifs associés.

La moyenne pondérée notée $\overline{x}$ de cette série est définie par :

$$\overline{x} = \frac{n_1 x_1 + n_2 x_2 + \cdots + n_p x_p}{n_1 + \cdots + n_p}$$

Exemple : capacité 1 p. 305 ou exercice 29 p 312

Propriété : Soitf_1, f_2, … , f_n les fréquences associées aux valeurs x_1, x_2, … x_n de la série statistique.

ALORS

$$\overline{x} = f_1 x_1 + f_2 x_2 + \cdots + f_p x_p$$

Démonstration :

Linéarité de la moyenne

Exemples : Dans une classe, la moyenne est de 9. Le professeur décide d'augmenter toutes les notes de 1 point. Quelle est la nouvelle moyenne ?

Dans une autre classe, la moyenne est de 10. Si le professeur augmente toutes les notes de 50%, quelle sera la nouvelle moyenne ?

Propriétés :

On considère une série statistique prenant comme valeurs $x_1, x_2, \dots x_p$. On note $\overline{x}$ sa moyenne.

Quel que soit le réel a , la série statistique prenant comme valeurs :

- $x_1 + a, x_2 + a, \dots x_p + a$ avec les mêmes fréquences a pour moyenne $\overline{x} + a$
- $ax_1, ax_2, \dots, ax_p$ avec les mêmes fréquences a pour moyenne $a\overline{x}$

Démonstration :

B. Variance et écart-type

Activité 2 : découverte de l'écart-type avec utilisation de Géogebra

Définition : On considère la série statistique $(x_1; n_1), (x_2; n_2), \dots, (x_p; n_p)$ où $x_1, x_2, \dots x_p$ sont les valeurs du caractère étudié et n_1,$n_2, \dots, n_p$ les effectifs associés. Soit $\overline{x}$ la moyenne de cette série.

L'écart type noté σ est définie par :

$$\sigma = \sqrt{\frac{n_1(x_1 - \overline{x})^2 + n_2(x_2 - \overline{x})^2 + \cdots + n_p(x_p - \overline{x})^2}{n_1 + n_2 + \cdots + n_p}}$$

L'écart-type mesure la dispersion des valeurs autour de la moyenne : plus il est grand, plus les valeurs sont dispersées et moins la moyenne représente de façon significative la série.

La variance se calcule ainsi : $\sigma = \sqrt{V}$

Exemple : capacité 2 p. 305/vérification à l'aide de la calculatrice.

Algorithme : Ecrire une fonction Python renvoyant la moyenne et l'écart type d'une série comportant 5 valeurs.

Exercices d'application :28, 40, 51.

II. Médiane, quartile et écart interquartile

Activité 3 et 4 p. 303 : découverte de la notion de quartile et d'écart interquartile.

A. Médiane :

Définition : Lorsqu'une série statistique est ordonnée, la **médiane** est la valeur qui partage cette série en deux séries de même effectif si l'effectif total est impair. Il y a donc autant de valeurs inférieures à la médiane que de valeurs supérieures.

Si l'effectif total N est pair, la médiane est la demi-somme des valeurs des termes de rang n et $n + 1$ dans cette série ordonnée.

Exemple : On considère la série statistique ordonnée constituée des valeurs suivantes : 7-7-8-9-10-10-11-11-11-12-14-15.

Déterminer la médiane de cette série.

Définition : L'effectif cumulé croissant de la valeur x_i est la somme des effectifs de toutes les valeurs du caractère inférieur ou égale à x_i.

Exemple :

B. Quartiles

Définition : Soit une série statistique dont les valeurs sont ordonnées par ordre croissant.

1) Le **premier quartile**, noté **Q_1**, est la plus petite valeur de la série telle qu'au moins 25% des valeurs sont inférieurs ou égales à ce nombre Q_1
2) Le **troisième quartile**, notée **Q_3**, est la plus petite valeur de la série telle qu'au moins 75% des valeurs sont inférieurs ou égales à ce nombre Q_3.

Méthode pour calculer les quartiles :

- Le rang du premier quartile d'une série d'effectif total N est le plus petit entier supérieur ou égal à $\frac{N}{4}$.
- Le rang du troisième quartile est le plus petit entier supérieur ou égale à $\frac{3N}{4}$

Exemple :

C. Ecart interquartile

Définition : On considère une série statistique de premier quartile Q_1 et de troisième quartile Q_3. L'écart interquartile de cette série est la différence

$$Q_3 - Q_1$$

L'intervalle [Q_1 ; Q_3] contient environ 50% des valeurs de la série. Il permet de mesurer la dispersion des valeurs d'une série statistique autour de la médiane.

Exemple : capacité 4 p. 307

Exercices d'applications : 54, 60.

Mode d'emploi calculatrice Numworks