SECONDE / COURS

Statistiques descriptives

Astrolab Cours · astrolabcours.fr

01 Le problème

Dans une petite entreprise, les salaires mensuels (en euros) sont : 1 800, 1 900, 1 900, 2 000, 2 100, 2 200 et 9 100 (le patron). Le salaire moyen est de 3 000 €. Est-ce qu'on peut dire que « on gagne bien sa vie dans cette entreprise » ?

Pour décrire une série, un seul nombre ne suffit pas. Il faut un indicateur de **position** et un indicateur de **dispersion**.

01 L'idée : la cagnotte et la file d'élèves

- La **moyenne**, c'est une **cagnotte partagée** : tout le monde met son salaire dans un pot, et on redistribue la même part à chacun. Un seul très gros salaire gonfle toutes les parts.
- La **médiane**, c'est l'élève **au milieu de la file**, quand tout le monde est rangé par taille. Peu importe que le dernier mesure 2 m ou 3 m, celui du milieu ne bouge pas.
- La **dispersion**, c'est la question : la file est-elle **serrée** (tout le monde à peu près pareil) ou **étalée** ?

01 Les indicateurs de position

Pour une série de valeurs $x_1, \dots, x_p$ d'effectifs $n_1, \dots, n_p$, d'effectif total N :

- **moyenne** : $\bar{x} = \dfrac{n_1 x_1 + n_2 x_2 + \cdots + n_p x_p}{N}$;
- **médiane** Me : la valeur qui partage la série **rangée** en deux groupes de même effectif. Si N est impair, c'est la valeur du milieu ; si N est pair, la moyenne des deux valeurs du milieu ;
- **premier quartile** Q_1 : la plus petite valeur telle qu'au moins 25 % des valeurs lui soient inférieures ou égales ; **troisième quartile** Q_3 : pareil avec 75 %.

Linéarité de la moyenne. Si on ajoute 50 € à tous les salaires, la moyenne augmente de 50 €. Si on les multiplie par 1,02, la moyenne est multipliée par 1,02.

Les indicateurs de dispersion

- **étendue** : valeur maximale moins valeur minimale ;
- **écart interquartile** : $Q_3 - Q_1$. La moitié centrale de la série est dans $[Q_1 ; Q_3]$;
- **écart-type** σ : il mesure l'écart moyen à la moyenne, $\sigma = \sqrt{\dfrac{n_1(x_1 - \bar{x})^2 + \cdots + n_p(x_p - \bar{x})^2}{N}}$.

© Astrolab Cours · astrolabcours.fr/maths/seconde/statistiques/ · Tous droits réservés : usage personnel uniquement, republication et usage commercial interdits

On associe la médiane avec l'écart interquartile (peu sensibles aux valeurs extrêmes), et la moyenne avec l'écart-type. En pratique, l'écart-type se calcule à la calculatrice ou au tableur.

Exemple guidé

Notes d'un groupe de 10 élèves : 8 (2 élèves), 10 (3 élèves), 12 (4 élèves), 15 (1 élève). Calcule la moyenne, la médiane et les quartiles.

1. **Moyenne** : $\bar{x} = \dfrac{2\times 8 + 3\times 10 + 4\times 12 + 1\times 15}{10} = \dfrac{109}{10} = 10{,}9$.
2. **Série rangée** : 8, 8, 10, 10, 10, 12, 12, 12, 12, 15.
3. **Médiane** : $N = 10$ est pair, on prend la moyenne des 5e et 6e valeurs, $Me = \dfrac{10+12}{2} = 11$.
4. **Quartiles** : $\frac{10}{4} = 2{,}5$, donc Q_1 est la 3e valeur, $Q_1 = 10$. $\frac{3\times 10}{4} = 7{,}5$, donc Q_3 est la 8e valeur, $Q_3 = 12$.
5. **Dispersion** : étendue $15 - 8 = 7$, écart interquartile $12 - 10 = 2$.

La moitié centrale du groupe a entre 10 et 12 : les notes sont plutôt serrées.

Retour au problème

La série rangée a 7 valeurs : la 4e est la médiane, $Me = 2\,000$ €. Le patron fait monter la moyenne à 3 000 €, alors que **6 salariés sur 7 gagnent au plus 2 200 €**.

Q_1 : $\frac{7}{4} = 1{,}75$, donc la 2e valeur, $Q_1 = 1\,900$. Q_3 : $\frac{3\times 7}{4} = 5{,}25$, donc la 6e valeur, $Q_3 = 2\,200$. L'écart interquartile est de 300 € : la file est serrée, sauf une valeur très à part. Ici, la médiane décrit bien mieux l'entreprise que la moyenne.

La moyenne et l'écart-type reviendront en Première avec l'espérance d'une variable aléatoire.

© Astrolab Cours · astrolabcours.fr/maths/seconde/statistiques/ · Tous droits réservés : usage personnel uniquement, republication et usage commercial interdits