

Chapitre 19 – Cours

Statistiques : dispersion et données regroupées

Dans ce chapitre, tu vas apprendre à :

- Calculer et interpréter une moyenne et un écart type.
- Exploiter des classes de même amplitude et des fréquences cumulées.
- Distinguer calcul exact et estimation sous une hypothèse de répartition.

Les mots pour comprendre

Définition – Un indicateur de dispersion

Un indicateur de dispersion renseigne sur l'étalement des valeurs autour du centre. Une même moyenne peut accompagner des séries très différentes.

Définition – L'écart-type

L'écart-type est la racine carrée de la variance. Il s'exprime dans la même unité que les données et vaut zéro pour une série constante.

1. Position et dispersion

Pour des valeurs x_i d'effectifs n_i , l'effectif total est $N = \sum n_i$ et la moyenne $\bar{x} = (\sum n_i x_i)/N$. Un indicateur de position décrit le centre d'une série ; il ne suffit pas à décrire sa dispersion.

Deux groupes peuvent avoir la même moyenne et des valeurs beaucoup plus éloignées de ce centre dans l'un que dans l'autre. La variance descriptive vaut $V = (\sum n_i (x_i - \bar{x})^2)/N$ et l'écart type $\sigma = \sqrt{V}$.

Cette convention utilise N , non $N - 1$. L'écart type s'exprime dans l'unité des données et mesure leur dispersion autour de la moyenne.

Il est nul si et seulement si toutes les valeurs sont identiques. Une formule équivalente est $V = (\sum n_i x_i^2)/N - \bar{x}^2$; un calcul machine peut toutefois produire une petite erreur d'arrondi.

2. Transformer les données

Ajouter une même constante b à toutes les valeurs ajoute b à la moyenne et laisse l'écart type inchangé.

Multiplier toutes les valeurs par a multiplie la moyenne par a et l'écart type par $|a|$. La valeur absolue est nécessaire puisque la dispersion ne peut pas devenir négative.

Ajouter ou supprimer une observation modifie à la fois la somme et l'effectif. On ne fait donc pas la moyenne de deux moyennes sans tenir compte de leurs effectifs.

Si un groupe de 10 élèves a une moyenne de 12 et un groupe de 20 une moyenne de 15, la moyenne commune est $(10 \times 12 + 20 \times 15)/30 = 14$.

3. Regrouper des données continues

Un tableau par classes conserve seulement le nombre d'observations dans chaque intervalle. Pour des classes de même amplitude, on construit un histogramme à rectangles contigus : une hauteur proportionnelle à l'effectif convient.

Les bornes des classes se lisent sur l'axe horizontal et l'unité de la hauteur est explicitée. Si les moyennes réelles de chaque classe sont connues, leur moyenne pondérée donne exactement la moyenne globale.

Si elles sont inconnues, remplacer chaque valeur par le centre de sa classe donne une estimation. Cette estimation est justifiée notamment sous une hypothèse de répartition uniforme dans chaque classe ; elle ne doit pas être annoncée comme une valeur exacte des données initiales.

4. Médiane et fréquences cumulées

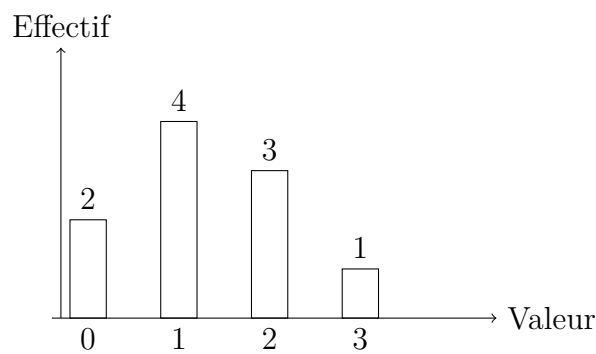
Les effectifs cumulés additionnent les effectifs jusqu'à une borne ; les fréquences cumulées divisent ces sommes par N .

Le seuil 50% permet de repérer une classe médiane. À partir des seuls effectifs par classes, la médiane exacte n'est généralement pas connue.

Sous une répartition uniforme dans la classe médiane $[a; b[$, on estime la médiane par interpolation linéaire. Si F_a est la fréquence cumulée à a et f celle de cette classe, l'estimation est $a + (b - a)(0,5 - F_a)/f$.

On indique l'hypothèse et on vérifie que le résultat appartient bien à la classe.

Voir la notion



Ce schéma donne un exemple concret de la notion. Les indications permettent de relier le dessin et les écritures mathématiques.

Deux exemples, pas à pas

Exemple résolu – 1 – Moyenne pondérée

Question. Les valeurs 5 et 9 ont des effectifs 3 et 1. Calculer la moyenne.

Résolution. $N = 4$ et $\bar{x} = (3 \times 5 + 9)/4 = 6$.

Exemple résolu – 2 – Suppression

Question. Une série de 10 valeurs de moyenne 12 contient la valeur 30. On la retire. Calculer la nouvelle moyenne.

Résolution. La somme devient $120 - 30 = 90$, pour 9 valeurs : la moyenne vaut 10.
