

1 Indicateurs de position ou de tendance centrale

Définition 1.

En statistique, un indicateur de position ou de tendance centrale est une valeur résumant une série statistique pour une variable quantitative. Les deux principaux sont la **moyenne** et la **médiane**. (Voir les définitions données dans les activités).

Remarque 1

La moyenne est sensible aux valeurs extrêmes alors que la médiane est beaucoup plus stable.

2 Indicateurs de dispersion

Définition 2.

En statistique, un indicateur de dispersion mesure la variabilité des valeurs d'une série statistique. Il est toujours positif et d'autant plus grand que les valeurs de la série sont étalées. Nous verrons en seconde l'**écart interquartile** et l'**écart-type**.

Ces indicateurs complètent l'information apportée par les indicateurs de position ou de tendance centrale, mesurée par la moyenne ou la médiane.

Définition 3.

Pour une série statistique de premier et troisième quartiles Q_1 et Q_3 :

- l'intervalle interquartile de la série est $[Q_1 ; Q_3]$;
- l'écart interquartile est $Q_3 - Q_1$.

Remarque 2

- *L'intervalle interquartile contient environ la moitié des valeurs de la série.*
- *L'écart interquartile est insensible aux valeurs extrêmes.*
- *L'écart interquartile est un **indicateur de dispersion** de la série : plus il est faible, plus la série est homogène.*

Définition 4.

La variance V (d'une série statistique) V est la moyenne des carrés des écarts à la moyenne.

L'**écart-type** σ d'une série statistique est $\sigma = \sqrt{V}$.

Dans la pratique, on utilisera la calculatrice pour déterminer l'écart-type.

Remarque 3

- L'écart-type mesure la **dispersion** d'une série statistique autour de la moyenne.
- La moyenne \bar{x} et l'écart-type σ s'expriment dans la même unité que les valeurs de la série, cela a un sens de considérer les intervalles $[\bar{x} - \sigma ; \bar{x} + \sigma]$, $[\bar{x} - 2\sigma ; \bar{x} + 2\sigma]$, etc. qui sont souvent utilisés en statistiques.

3 Résumé d'une série statistique

Méthode.

On peut résumer une série statistique, c'est-à-dire en donner une tendance globale, par :

- le couple médiane-écart interquartile, qui n'est pas sensible aux valeurs extrêmes : on le privilégie donc quand on étudie une série dont les valeurs extrêmes sont moins *importantes* ou moins *significatives* que les valeurs *centrales*;
- le couple moyenne-écart-type, qui est sensible aux valeurs extrêmes : on le privilégie donc quand on étudie une série dont les valeurs extrêmes sont aussi *importantes* ou aussi *significatives* que les autres.

Dans les deux cas, on utilise un indicateur de position (la médiane ou la moyenne) et un indicateur de dispersion (l'écart interquartile ou l'écart-type).

Exemple.

Pour chacune des deux situations suivantes, est-il est préférable de résumer la série statistique correspondante par le couple médiane-écart interquartile ou par le couple moyenne-écart-type?

- **Situation 1** : On étudie la série statistique des salaires et allocations chômage des Français en 2014 en vue d'en observer les inégalités.
- **Situation 2** : On étudie les résultats d'une enquête d'un fabricant de chaussures portant sur la taille de chaussure de ses clients afin de déterminer la production de quelles pointures privilégier.

Réponses :

- Dans la situation 1, les valeurs extrêmes sont très importantes puisque ce sont elles qui illustrent les plus grandes inégalités : on préférera donc le couple moyenne-écart-type.
- Dans la situation 2, le fabricant souhaite savoir quelles pointures sont les plus portées et ne s'intéresse donc pas aux très petites et très grandes pointures peu portées par ses clients mais plutôt aux valeurs centrales : on préférera donc le couple médiane-écart interquartile.