# STATISTIQUES

## I Série et indicateurs

**À préciser :** population, caractère (quantitatif discret / continu, ou qualitatif), unité, source.
$N = \sum n_i$, $f_i = n_i/N$, $\sum f_i = 1$.

| Indicateur | Calcul / lecture | Sens |
|---|---|---|
| Moyenne $\overline{x}$ | $\overline{x} = \dfrac{\sum n_i x_i}{N} = \sum f_i x_i$ | Centre, sensible aux extrêmes. |
| Médiane | Données triées ; valeur centrale si N impair, moyenne des deux centrales si N pair. | Partage la série en deux moitiés. |
| $Q_1$,$Q_3$ | Rangs $\lceil N/4 \rceil$, $\lceil 3N/4 \rceil$. | Écart interquartile $Q_3 - Q_1$. |
| Étendue | Maximum − minimum. | Dispersion totale. |
| Écart type σ | $\sqrt{\dfrac{\sum n_i (x_i - \overline{x})^2}{N}}$ | Dispersion autour de la moyenne ; même unité. |

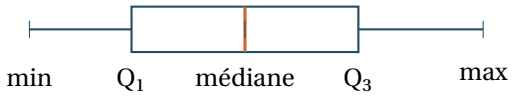
## II Transformer ou réunir des données

$$\overline{ax+b} = a\overline{x} + b, \qquad \overline{x}_{\text{ajout}} = \frac{N\overline{x} + y}{N+1}, \qquad \overline{x}_{\text{retrait}} = \frac{N\overline{x} - y}{N-1} \quad (N > 1).$$

**Réunir des groupes :** $\overline{x} = (N_1\overline{x}_1 + \cdots + N_p\overline{x}_p)/(N_1 + \cdots + N_p)$. Pour la médiane : recalculer les rangs après ajout / retrait.

## III Représenter et estimer avec des classes

**Histogramme :** aire du rectangle proportionnelle à l'effectif ou la fréquence ; pas la seule hauteur si amplitudes différentes.
**Cumulées :** sommes progressives ; polygone des bornes de classe, du niveau 0 au niveau 1.
**Hypothèse d'uniformité :** centres des classes pour estimer la moyenne ; interpolation au niveau 0,5 pour estimer la médiane.
**Sans uniformité :** localiser les valeurs centrales ; si un cumul vaut exactement 50 %, elles encadrent une borne de classe et leur moyenne reste inconnue.



Boîte schématique : cinq repères.

## IV Tableaux croisés et comparaison

| Lecture | Dénominateur |
|---|---|
| Fréquence marginale | Total de ligne ou de colonne / effectif total. |
| Fréquence conditionnelle | Case commune / total du groupe imposé par « parmi ». |

**Comparer :** centre *et* dispersion, mêmes unités, avec le graphique et le contexte.
$(\overline{x}\,;\sigma)$ : sensible aux extrêmes ; (médiane ; $Q_3 - Q_1$) : moins sensible.
**Tableur :** `ET` : deux critères ; `OU` : au moins l'un ; `NON` : contraire.