2 Indicateurs d'une série statistique

2.1 Moyenne

Définition (Moyenne)**.** La *moyenne* d'une série de caractères $x_1, x_2, \ldots x$
et d'effectifs $n_1, n_2 \ldots n_k$ est le nombre réel notée $\bar{x}$ valant :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \ldots + n_kx_k}{n_1 + n_2 + \ldots + n_k}.$$

Exemple 1. Calculer la moyenne des salaires de l'entreprise suivante.

Salaire	1450	1500	1550	1600	1650	1700	1750
Effectif	5	2	5	7	3	6	7

La moyenne des salaires est : $\frac{5\times1450+2\times1500+5\times1550+\cdots+7\times1750}{5+2+5+\cdots+7}$, soit environ 1617,14€.

2.2 Médiane

Définition (Médiane)**.** La *médiane* d'une série statistique est un nombre m tel que la moitié des effectifs étudiés ait une valeur inférieure ou égale à m, et l'autre moité ait une valeur supérieure ou égale.

Méthode (Calcul de la médiane)**.** Soit une série statistique de taille N, triée par ordre croissant.

- Si N est impair, la médiane est la valeur de rang $\frac{N+1}{2}$.
- Si N est pair, la médiane est la moyenne des valeurs de rang $\frac{N}{2}$ et $\frac{N}{2} + 1$.

Remarque. La moyenne et la médiane sont des indicateurs de *position* (ils indiquent si les valeurs de la série sont, dans l'ensemble, plutôt grandes ou plutôt petite).

2.3 Quartiles

Définition (Quartiles)**.**

— Le *premier quartile* d'une série (noté Q_1) est la plus petite valeur telle qu'au moins 25% des valeurs lui soit inférieure ou égale.

— Le *troisième quartile* d'une série (noté Q_3) est la plus petite valeur telle qu'au moins 75% des valeurs lui soit inférieure ou égale.

— *L'écart interquartile* est le nombre réel $Q_3 - Q_1$.

Méthode. On trie les effectifs par ordre croissant, et on applique la définition.

Exemple 2. Calculer la médiane et les quartiles des salaires de l'entreprise suivante.

Salaire	1450	1500	1550	1600	1650	1700	1750
Effectif	5	2	5	7	3	6	7

On ajoute une ligne pour calculer les effectifs cumulés croissants (notés ECC). Remarque : cela fonctionnerait aussi si on avait les fréquences au lieu des effectifs.

Salaire	1450	1500	1550	1600	1650	1700	1750
Effectif	5	2	5	7	3	6	7
ECC	5	7	12	19	22	28	35

Il y a 35 salaires.

— Puisque $\frac{35}{4} = 8,75$, le premier quartile est le premier salaire dont l'ECC dépasse 8,75, c'est-à-dire $Q_1 = 1550$.

— Puisque $\frac{3\times35}{4} = 26,25$, le troisième quartile est le premier salaire dont l'ECC dépasse 26,25, c'est-à-dire $Q_3 = 1700$.

Remarque. Les quartiles (et l'espace interquartiles en particulier) sont des indicateurs de *dispersion* (ils indiquent si les valeurs de la série sont toutes proches les unes des autres, au étalées (certaines valeurs très grandes, d'autres très petites)).

Remarque. La médiane et les quartiles coupent la série statistique en quatre séries de taille à peu près égales.

Par exemple, si on étudie les scores obtenus par les équipes d'une compétition sportive, il y aura :

— environ un quart des équipes ayant obtenu un score inférieur au premier quartile Q_1 ;

— environ un quart des équipes ayant obtenu un score compris entre le premier quartile Q_1 et la médiane ;

— environ un quart des équipes ayant obtenu un score compris entre la médiane et le troisième quartile Q_3 ;

— environ un quart des équipes ayant obtenu un score supérieur au troisième quartile Q_3.