Chapitre 1

Statistiques

Activité qu'est-ce qu'une série statistique ?

A Quartiles

Nous nous intéressons ici aux séries statistiques dont l'étude porte sur des **caractères** prenant des valeurs numériques qui peuvent naturellement être ordonnées.

Exemple des notes, des prix, des longueurs, ...

1 Définition

Rappel La médiane est la valeur du caractère qui sépare la population en deux groupes de même taille : celui pour lequel le caractère est inférieur et celui pour lequel le caractère est supérieur.
Autre manière de voir : lorsque l'on calcule les effectifs cumulés croissants, la médiane est la plus petite valeur du caractère qui fait atteindre ou dépasser la moitié de la population.
Certaines définitions diffèrent, pour savoir si la médiane doit être une valeur atteinte ou non. Dans le cas où le caractère est regroupé par classe, on parle parfois de classe médiane.
On note parfois M_e la médiane.

Proposition Soit $(x_1, x_2, \ldots, x_N)$ une série statistique ayant une médiane M_e. On appelle dispersion des écarts absolus le nombre défini pour tout réel x par :

$$f(x) = |x - x_1| + |x - x_2| + \cdots + |x - x_N| = \sum_{i=1}^{k} |x - x_i|$$

La fonction f admet un minimum atteint lorsque $x = M_e$.

Preuve : voir page 300. □

Définition Lorsque l'on calcule les effectifs cumulés croissants,
- Le premier quartile est la (plus petite) valeur du caractère qui fait atteindre ou dépasser le quart de la population.
- Le troisième quartile est la (plus petite) valeur du caractère qui fait atteindre ou dépasser les trois quarts de la population.
- On note Q_1 le premier quartile et Q_3 le troisième quartile.
- L'intervalle interquartile est $[Q_1 ; Q_3]$. L'écart interquartile est le nombre $I = Q_3 - Q1$.

Remarque L'intervalle interquartile contient donc environ la moitié de l'effectif global par définition.

Remarque La médiane peut être considérée comme le deuxième quartile.

Remarque On peut définir de manière similaire les déciles (le premier et le neuvième), qui s'intéressent aux dixièmes de la population.
→ **Exercice** exercice résolu pages 304-5 (sauf dernière question, pour sections d'après).
→ **Exercices** 33,34,35p316

2 Diagramme en boîte

Lire : 4p298 et page 299.
→ **Exercice** exercice résolu 5 page 306 (revoir page 305)
→ **Exercices** 36,38p316 et 41p318 (sauf 2.)

3 Influence d'une transformation affine

Faire subir une transformation affine à une série statistique, c'est appliquer une fonction affine à l'ensemble des valeurs de la série.

Exemple Soit $(x_1, x_2, \ldots, x_N)$ une série statistique. Soit $f : x \mapsto ax + b$ une fonction affine. Si l'on fait subir une transformation affine à la série statistique avec la fonction f, on obtient la série statistique $(ax_1 + b, ax_2 + b, \ldots, ax_N + b)$.

Proposition On suppose $a > 0$. Soit M_e la médiane et I l'écart interquartile de la série $(x_1, x_2, \ldots, x_N)$. La médiane et l'écart interquartile de la série statistique $(ax_1 + b, ax_2 + b, \ldots, ax_N + b)$ sont alors respectivement $aM_e + b$ et aI.

Preuve : La médiane et les quartiles sont des valeurs de la série statistique. Si l'on suppose la série rangée par ordre croissant, on a :
- $M_e = x_e$ tel que e est le plus petit indice i tel que $i \geq N/2$;
- $Q_1 = x_\alpha$ tel que α est le plus petit indice i tel que $i \geq N/4$;
- $Q_3 = x_\gamma$ tel que γ est le plus petit indice i tel que $i \geq 3N/4$.

La série transformée est toujours rangée par ordre croissant car f est croissante, a étant positif. Les indices conservant leur propriété donnée ci-dessus, la nouvelle médiane est $ax_e + b = aM_e + b$ et le nouvel écart interquartile est $ax_\gamma + b - (ax_\alpha + b) = a(x_\gamma - x_\alpha) = aI$. □
→ **Exercices** 40p317 (conversion de monnaie)