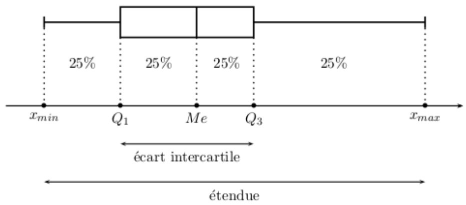
Fiche méthode n°10 – Statistiques

Vocabulaire :

- *Les modalités* d'une série statistiques sont les valeurs qu'elle prend. On les note $x_1, x_2, \ldots$
- *L'effectif d'une modalité* est le nombre de valeurs égales à cette modalité. On les note $n_1, n_2, \ldots$ et n : *effectif total.*
- *L'effectif cumulé* d'une modalité est le nombre de valeurs qui lui sont inférieures ou égale.
- *La fréquence* d'une modalité est la proportion de valeurs égales à cette modalité. On les note $f_1, f_2, \ldots$
- *La fréquence cumulée* d'une modalité est la proportion de valeurs qui lui sont inférieures ou égale.
- *La fonction de répartition empirique* est définie par : F(x) = (proportion de valeurs de la série inférieures ou égales à x)

Diagrammes :

- *En bâtons* : (abscisses = modalités) et (hauteurs des bâtons = effectifs)
- *En boîte* : Cf ci-contre



- *FCC* : On place les points avec (abscisses = modalités) et (ordonnées = fréquances cumulées)

  On créée une fonction en escalier en prolongeant vers la droite par des segments horizontaux.

Pour calculer la moyenne $\overline{x}$ d'une série statistique :

- On divise la somme des valeurs par le nombre de valeurs.
- On utilise l'une des formules $\overline{x}=\frac{1}{n}\sum_i n_i x_i=\sum_i f_i x_i$ avec le tableau des effectifs/fréquences.
- Si x est obtenue par transformation affine d'une autre série u : x = a u + b , alors on a $\overline{x}=a\overline{u}+b$

Pour calculer la médiane Me(x) d'une série statistique :

- Si le nombre de valeurs est impair, la médiane Me est la valeur centrale.

  Si le nombre de valeurs est pair, la médiane est la moyenne des deux valeurs centrales
- Si x est obtenue par transformation affine d'une autre série u : x = a u + b , alors on a $Me(x)=a\,Me(u)+b$

Pour calculer le premier quartile Q_1 d'une série statistique :

- On cherche la première modalité dont la fréquence cumulée dépasse 25%
- On calcule un quart de l'effectif total ($\frac{1}{4}\times n$) , et on arrondit à l'entier supérieur.
  - Le résultat obtenu est le rang du premier quartile dans la liste triée par ordre croissant.
- Graphiquement, on regarde l'abscisse du premier point de la courbe des FCC dont l'abscisse est au-dessus de 0,25.

Adapter cet argument pour le troisième quartile, les déciles, ou généralement un quantile d'ordre α

Pour calculer la variance σ_x^2 d'une série statistique :

- On utilise une des formules de la définition : $\sigma_x^2=\frac{1}{n}\sum_i n_i(x_i-\overline{x})^2=\sum_i f_i(x_i-\overline{x})^2$
- On utilise la *formule de Koenig* : $\sigma_x^2=\overline{x^2}-\overline{x}^2$
- Si x est obtenue par transformation affine d'une autre série u : x = a u + b , alors on a $\sigma_x^2=a^2\sigma_u^2$.