

Chap V Statistique descriptive

- Représentations graphiques d'une série,
- Effectifs, fréquences, cumul,
- Moyenne, médiane, quartiles,
- Utilisation de logiciels et de la calculatrice.

0 Intro.

Faire les 6 exos p.164 + les 2 activités p166-167

I Présentation d'une série statistique.

A la maison, faire recopier la p.170 et étudier n°1 et 2 p.171

1 Effectifs et fréquences

Définitions

Dans une série statistique,

- l'**effectif** d'une valeur est le nombre de données correspondant à cette valeur ;
- la **fréquence** d'une valeur est $f = \frac{\text{effectif de la valeur}}{\text{effectif total}}$.

EXEMPLE

En lançant 10 fois un dé, on obtient : 2 ; 4 ; 6 ; 6 ; 3 ; 4 ; 4 ; 5 ; 3 ; 6.

L'effectif total est 10.

La valeur 6 apparaît 3 fois : son effectif est 3, sa fréquence est $\frac{3}{10} = 0,3$.

Valeur x_i	2	3	4	5	6
Effectif n_i	1	2	3	1	3
Fréquence f_i	0,1	0,2	0,3	0,1	0,3

Méthode

Quand on cherche le nombre de valeurs de la série inférieures ou égales à une valeur donnée, on est amené à ajouter (cumuler) les effectifs :

Valeur x_i	2	3	4	5	6
Effectif n_i	1	2	3	1	3
Effectif cumulé	1	1 + 2 = 3	1 + 2 + 3 = 6	1 + 2 + 3 + 1 = 7	1 + 2 + 3 + 1 + 3 = 10

Il y a 6 valeurs de la série inférieures ou égales à 4

(1 valeur égale à 2 + 2 valeurs égales à 3 + 3 valeurs égales à 4).

On retrouve l'effectif total

En pratique, on cumule pas à pas les effectifs (voir exercice résolu 1).

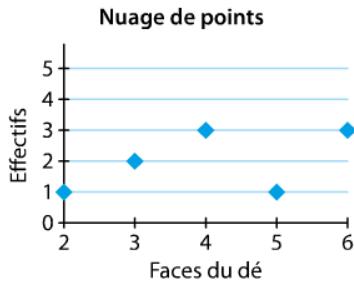
On peut cumuler de la même façon des fréquences (voir exercice résolu 2).

2 Graphiques

Le choix d'un graphique dépend du type de série et de ce que l'on veut montrer :

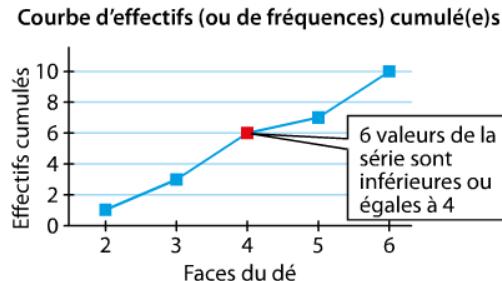
- un diagramme en bâtons (pour des valeurs discrètes ou qualitatives) et un histogramme (pour des valeurs numériques regroupées en classes) représentent les effectifs ou fréquences pour chaque valeur ou classe ;
- un diagramme circulaire montre la part relative de chaque valeur ou classe.

On utilisera aussi les graphiques suivants (relatifs à l'exemple ci-dessus).



On peut aussi relier les points par des segments.

Un nuage de points peut être utilisé pour montrer l'évolution d'une série.



Les valeurs de la série sont en abscisse, les effectifs (ou fréquences) cumulé(e)s en ordonnée.

On relie les points par des segments.

II Paramètres de position (tendance centrale)

Acti 3 p.168 puis exo complet (fiche)

Étudier exos corrigés n°1-2-3-5-7 p.171 à 175

1°) La moyenne.

Elle se calcule avec les effectifs ou les fréquences.

Soit une variable statistique d'effectif N prenant p valeurs ($p \leq N$).

Valeurs	x_1	x_2	x_p
Effectif	n_1	n_2	n_p
Fréquence	$f_1 = \frac{n_1}{N}$	$f_2 = \frac{n_2}{N}$	$f_p = \frac{n_p}{N}$

$$N = n_1 + n_2 + n_3 + \dots + n_p$$

$$f_1 + f_2 + f_3 + \dots + f_p = 1$$

Définition : La moyenne de la série est le nombre, notée $\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N}$

$$\text{ou } \bar{x} = f_1 x_1 + f_2 x_2 + \dots + f_p x_p.$$

2°) Une médiane.

C'est une valeur notée Me . Elle partage l'effectif total en 2 sous-groupes de même taille : ceux qui ont une valeur $\leq Me$ et ceux qui ont une valeur $\geq Me$.

Méthode : La liste des N données est rangée dans l'ordre croissant.

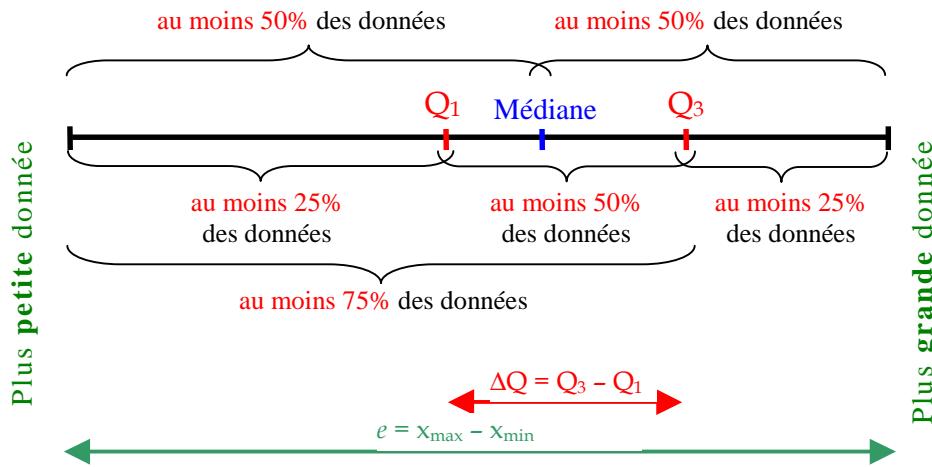
- Si la série est de taille **impaire**, la médiane est la donnée de rang $\frac{N+1}{2}$.
- Si la série est de taille **paire**, toute valeur comprise entre des données de rang $\frac{N}{2}$ et $\frac{N}{2} + 1$ peut servir de médiane (*et pas nécessairement leur moyenne*).

3°) Les quartiles.

Définition : La liste des N données est rangées dans l'ordre croissant.

- Le **premier quartile** est noté Q_1 . C'est une des valeurs de la liste et elle réalise le partage suivant : au moins $\frac{1}{4}$ des valeurs doivent être inférieures ou égales à Q_1 . C'est bien entendu la plus petite valeur réalisant cela.
- Le **troisième quartile** est noté Q_3 . C'est une des valeurs de la liste et elle réalise le partage suivant : au moins $\frac{3}{4}$ des valeurs doivent être inférieures ou égales à Q_3 . C'est bien entendu la plus petite valeur réalisant cela.

NB : Pour obtenir une médiane et les quartiles, il faudra souvent calculer les ECC.



III Avec la calculatrice.

Voir n°6 p.174 Tableurs et calculatrices donnent parfois des quartiles faux :((

IV Paramètre de dispersion.

Voir n°4 p.174 + Évidemment p.188-189 QCM, V/F, évaluer ses capacités

Définitions : • $e = x_{\max} - x_{\min}$ c'est l'étendue de la série.

- L'amplitude d'une série est l'intervalle $[x_{\min} ; x_{\max}]$.
- L'écart interquartile est le nombre $\Delta Q = Q_3 - Q_1$

Remarque 1 : ΔQ est un bon indicateur de dispersion car est moins sensible aux valeurs extrêmes que e

Remarque 2 : L'intervalle $[Q_1 ; Q_3]$ donne un intervalle centré sur la médiane qui contient env. 50% des données.