

STATISTIQUE- COURS

Vocabulaire :

Étude statistique : une étude statistique se fait sur une **population** (élèves d'une classe, habitants d'une ville, les animaux d'un zoo, les voitures fabriquées par une usine...). Elle consiste à étudier un **caractère** qui peut être **quantitatif** (avec des nombres : salaire mensuel, poids, prix, ...) ou **qualitatif** (sans nombre : la couleur des yeux, la forme des lunettes,...).

Caractère : il présente plusieurs valeurs notées x_i . Lorsque le nombre de valeurs possibles est trop grand (ex : la taille d'une personne), on regroupe plusieurs valeurs dans une **classe** (un intervalle)

Effectif : nombres de fois n_i qu'apparaît le caractère x_i (ex : 3 personnes ont les yeux bleus).

Fréquence : c'est le rapport entre l'effectif d'une valeur i et l'effectif total N de la série.

$$f_i = \frac{\text{effectif d'une valeur}}{\text{effectif total}}$$

Indicateurs de tendance centrale :

La moyenne : Pour p valeurs de caractères différentes

$$\bar{x} = \frac{1}{N} \times \sum n_i \times x_i = \frac{n_1 \times x_1 + n_2 \times x_2 + \dots + n_p \times x_p}{N}$$

En fréquence : $\bar{x} = \sum f_i \times x_i = f_1 \times x_1 + f_2 \times x_2 + \dots + f_p \times x_p$

Interprétation : En approximant, on peut considérer que la valeur du caractère étudié est égale à \bar{x} pour chaque effectif constituant la population. Pour une série statistique définie par des classes, on prend pour valeurs des x_i le milieu de chaque classe.

Linéarité de la moyenne : Si on multiplie/ajoute chaque valeur du caractère par un même nombre alors la nouvelle moyenne est obtenue en ajoutant/multipliant l'ancienne moyenne par ce nombre.

Si $y_i = a \times x_i + b$ alors $\bar{y} = a \times \bar{x} + b$

La médiane : La moitié de la population a un caractère inférieur à la médiane et l'autre moitié un caractère supérieur. On la note en général, **Me**. Pour la déterminer, il faut ranger les valeurs du caractère par ordre croissant (les regrouper dans un tableau si trop nombreux) :

- Si l'effectif total N , est impair alors la médiane est la $\frac{N+1}{2}$ ème valeur.

- Si l'effectif total N , est pair alors la médiane est la moyenne entre la $\frac{N}{2}$ ème valeur et la $\frac{N}{2} + 1$ ème valeur.

Le 1^{er} quartile : Au moins 25 % de la population a un caractère inférieur au premier quartile. On le note en général, **Q₁**. Pour le déterminer, il faut ranger les valeurs du caractère par ordre croissant.

Le premier quartile est alors la $\frac{N}{4}$ ème (arrondi à l'entier supérieur) valeur.

Le 3^{ème} quartile : Au moins 75 % de la population a un caractère inférieur au premier quartile. On le note en général, **Q₃**. Pour le déterminer, il faut ranger les valeurs du caractère par ordre croissant.

Le troisième quartile est alors la $\frac{3N}{4}$ ème (arrondi à l'entier supérieur) valeur.

Indicateurs de dispersion :

Ecart-type et Variance :

Variance : $V = \frac{1}{N} \sum n_i (x_i - \bar{x})^2 = \frac{1}{N} (n_1 (x_1 - \bar{x})^2 + n_2 (x_2 - \bar{x})^2 + \dots + n_p (x_p - \bar{x})^2)$

Qui peut aussi se calculer : $V = \frac{1}{N} (\sum x_i^2) - \bar{x}^2$

En fréquence : $V = \sum f_i (x_i - \bar{x})^2$

Ecart-type : $\sigma = \sqrt{V}$

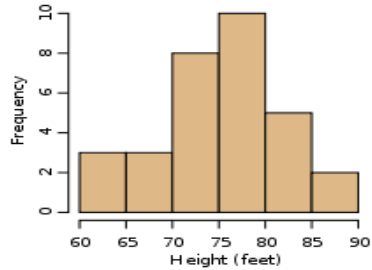
Interprétation : Plus l'écart-type est grand, plus les valeurs sont dispersées autour de la moyenne (grands écarts de valeurs). La série est d'autant plus homogène que l'écart-type est petit (sinon hétérogène). Environ 95 % des valeurs sont comprises dans l'intervalle : $[\bar{x} - 2\sigma; \bar{x} + 2\sigma]$. Pour décrire une série on peut utiliser le couple $(\bar{x}; \sigma)$

Ecart interquartile : $Q_3 - Q_1$

Interprétation : Plus l'écart interquartile est grand, plus les valeurs sont dispersées autour de la médiane. Entre Q_1 et Q_3 , il y a 50 % des valeurs de la série.

Pour décrire une série on peut utiliser le couple $(Me; Q_3 - Q_1)$

Représentations graphiques :



Histogramme : pour les séries avec des classes.

Fréquence ou effectif en ordonnée et caractère en abscisse. Caractère qualitatif ou quantitatif. Aire des rectangles proportionnelle aux effectifs.

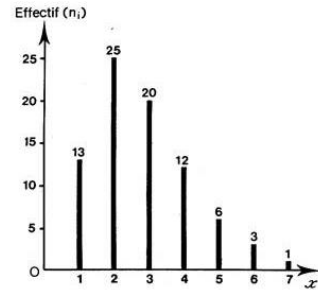
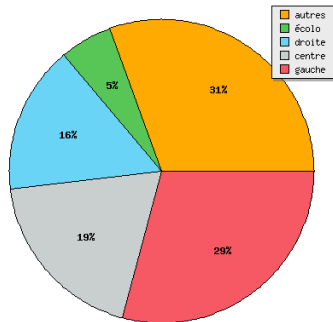


Diagramme en bâtons : pour les séries sans classe.

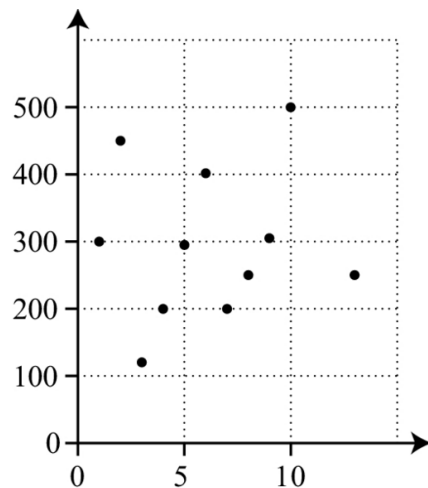
Fréquence ou effectif en ordonnée et caractère en abscisse. Caractère qualitatif ou quantitatif. Hauteur des bâtons proportionnelle aux effectifs.

Couverture web des forces politiques
Mesures entre le 25/10 et 31/10



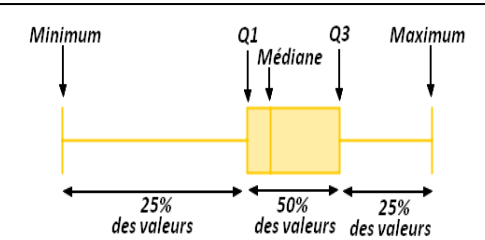
Camembert : représentation circulaire avec des fréquences converties en pourcentages.

Caractère qualitatif ou quantitatif avec classes. L'angle au centre de chaque secteur est proportionnel au pourcentage $100\% = 360^\circ$

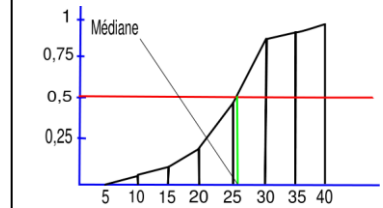


Nuage de point : pour établir une relation entre les effectifs et les valeurs.

Fréquence ou effectif en ordonnée et caractère en abscisse (sans classe). Chaque point a pour coordonnées (valeur, effectif). Caractère quantitatif



Boîte à moustaches : Permet d'observer la répartition des valeurs par rapport à la médiane et aux quartiles. Caractère quantitatif sans classe.



Courbe des effectifs (ou fréquences) cumulé croissant : Permet de trouver graphiquement la médiane et les quartiles des caractères quantitatifs avec classes. (Dans l'exemple ci-dessus, fréquence en ordonné et valeurs du caractère en abscisse)

Méthodes (exercices) :

	<u>Hachette</u>	<u>Hatier</u>	<u>Mes exos</u>	<u>Sesamaths</u>	<u>Mathx</u>
A) Calculer une moyenne	32-35,37,47	31,30,33,38	Ex. 1	179,180	26-33
B) Calculer un écart-type	36	35	Ex. 2	181	44,47
C) Calculer médiane	39	46,47	Ex. 3		34-38
D) Calculs de quartiles et écart-interquartile	39,41	46,47,49	Ex. 4	181	43,44,45
E) Interpréter	43,46,52	25,26	Ex. 5		39-42,48-57

Exercices de synthèse :

	<u>Hachette</u>	<u>Hatier</u>	<u>Mes exos</u>	<u>Sesamaths</u>	<u>Mathx</u>
Algorithmes	30	23,39	Ex. 7		60,64
Synthèse	48-50,57,58,65-68	22,58-68	Ex. 6	182-188	58-59,61
QCM					74-78
Vrai/faux	38	69			67-70

