

STATISTIQUES

La statistique est une science qui se propose de décrire, de représenter, d'analyser et d'interpréter des phénomènes collectifs susceptibles d'être quantifiés.

Depuis l'Antiquité jusqu'au Moyen-âge, la statistique avait surtout pour objectif de recenser la population afin de renseigner les États (même racine que " statistique " : Status, en latin). Le vocabulaire employé en statistique en découle (population, individus ...). De nos jours, les états modernes possèdent des organismes chargés de mesurer l'activité économique sous ses différents aspects. En France, c'est l'Institut national de la statistique et des études économiques (INSEE, site internet : www.insee.fr).

Depuis plusieurs décennies, la statistique a évolué: elle a pris une ampleur considérable dans les secteurs publics et privés qui possèdent leurs propres services de statistique pour mesurer, prévoir et prendre des décisions.

La démarche statistique commence toujours par une question que l'on se pose. C'est à partir de cette question qu'est élaborée l'enquête.

I. Rappel sur le vocabulaire :

Réaliser une étude statistique consiste à classer les **individus** d'une **population** en fonction d'un **caractère** (ou **variable**). Les premières études statistiques étaient démographique, on en a gardé le vocabulaire.

Exemple :

Classer les **élèves** d'une **classe** en fonction de leur **âge**.

Classer les **voitures** garées sur un **parking** en fonction de leur **couleur**.

Classer les **joueurs** d'une **équipe** de foot en fonction de leur **poste**.

Classer des **forfaits** d'un **opérateur** téléphonique en fonction de leur **prix**.

Le caractère étudié peut être **quantitatif** (lorsqu'il prend des valeurs numériques : âge, prix...) ou **qualitatif** (dans les autres cas : couleur, poste des joueurs de foot...).

Exemple : On a relevé l'âge des élèves d'une classe de 35 élèves :

Âge (caractère)	14	15	16	17	TOTAL
Effectif	1	27	5	2	35
Effectif cumulé croissant	1	28	33	35	
Effectif cumulé décroissant	35	34	7	2	
Fréquence (%)	0,029 (2,9%)	0,771 (77,1%)	0,143 (14,3%)	0,057 (5,7%)	1 (100%)
Fréquence croissante cumulée	0,029 (2,9%)	0,8 (80%)	0,943 (94,3%)	1 (100%)	
Fréquence décroissante cumulée	1 (100%)	0,971 (97,1%)	0,2 (20%)	0,057 (5,7%)	

Pour chaque valeur du caractère, on peut indiquer l'**effectif** (le nombre d'individus) ou la **fréquence** (la proportion d'individu par rapport à la totalité de la population).

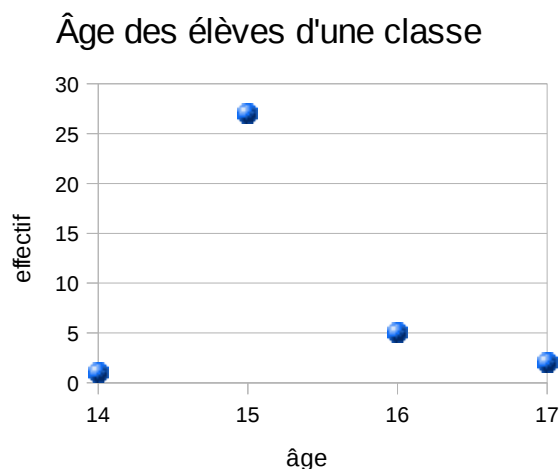
Cette fréquence peut s'exprimer sous la forme :

- d'un nombre décimal entre 0 et 1 (Exemple : 0,057)
- d'un pourcentage (Exemple : 5,7%)
- d'une fraction (Exemple : $\frac{57}{100}$)

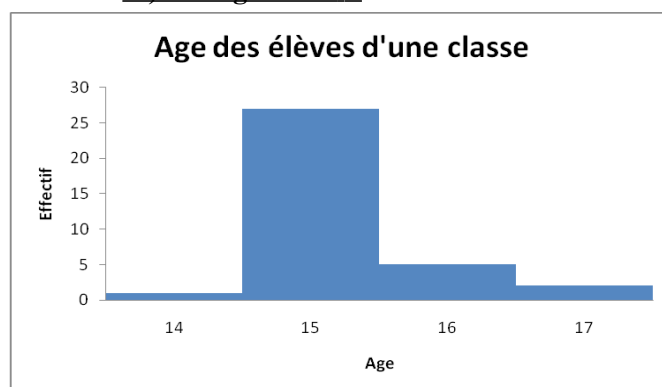
II. Exemples de représentations graphiques :

Dans ce paragraphe, nous reprenons les données du paragraphe précédent sur l'âge des élèves d'une classe de 35 élèves.

1°) Nuage de points :



2°) Histogramme :



Remarque : dans un histogramme, c'est l'aire de la « barre » qui doit être proportionnelle à l'effectif, et non pas sa hauteur. Il faut donc faire très attention lorsque le pas n'est pas constant.

L'histogramme est bien adapté aux caractères quantitatifs dont les valeurs sont regroupées en classes.

Remarque :

Les deux représentations graphiques ci-dessus ne sont que deux exemples parmi bien d'autres. Vous en connaissez d'autres comme le diagramme en bâtons, le diagramme circulaire ou semi-circulaire, ou encore la courbe.

III. Médiane et quartiles :

1°) Médiane :

Définition :

C'est la valeur **Me** pour laquelle la **moitié au moins** des individus de la population a une valeur pour le caractère étudié **supérieure ou égale** à la médiane et la **moitié au moins** a une valeur **inférieure ou égale** à la médiane.

Exemples :

- Si N est impair, la médiane est sa valeur centrale :

Notes de 11 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; **11** ; 12 ; 12 ; 13 ; 14 ; 15. La médiane est la 6^{ème} valeur, soit Me = 11

- Si N est pair, la médiane est la demi-somme (moyenne) des deux valeurs centrales :

Notes de 10 élèves : 5 ; 6 ; 7 ; 7 ; **9 ; 12** ; 12 ; 13 ; 14 ; 15. La médiane est la moyenne entre la 5^{ème} et la 6^{ème} valeur, soit $Me = \frac{9+12}{2} = 10,5$.

2°) Quartiles :

Définitions :

On considère une série discrète dont les valeurs sont rangées dans l'ordre croissant.

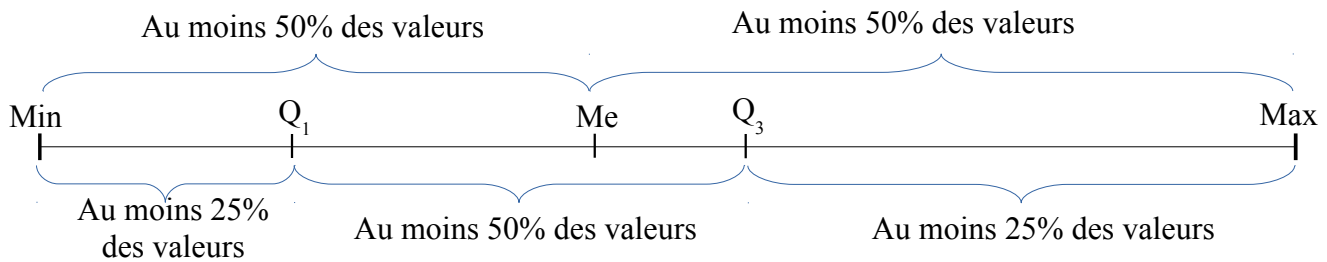
- Le **premier quartile** d'une série, noté **Q₁**, est la plus petite valeur de la série pour laquelle au moins **un quart (25%)** des données sont inférieures ou égales à **Q₁**.

- Le **troisième quartile** d'une série, noté **Q₃**, est la plus petite valeur de la série pour laquelle au moins **trois quarts (75%)** des données sont inférieures ou égales à **Q₃**.

Définitions :

On appelle intervalle interquartile, l'intervalle $[Q_1 ; Q_3]$.

On appelle l'écart interquartile, noté Q , l'amplitude de l'intervalle $[Q_1 ; Q_3]$ c'est à dire le nombre $Q = Q_3 - Q_1$.



Exemple 1 :

Caractère quantitatif discret (retour aux notes des élèves)

- Notes de 11 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 11 ; 12 ; 12 ; 13 ; 14 ; 15.

Me=11 $Q_1 = 7$ et $Q_3 = 13$

- Notes de 10 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 12 ; 12 ; 13 ; 14 ; 15.

Me=10,5 $Q_1 = 7$ et $Q_3 = 13$

- Notes de 12 élèves : 5 ; 6 ; 7 ; 7 ; 9 ; 11 ; 12 ; 12 ; 13 ; 14 ; 15 ; 16

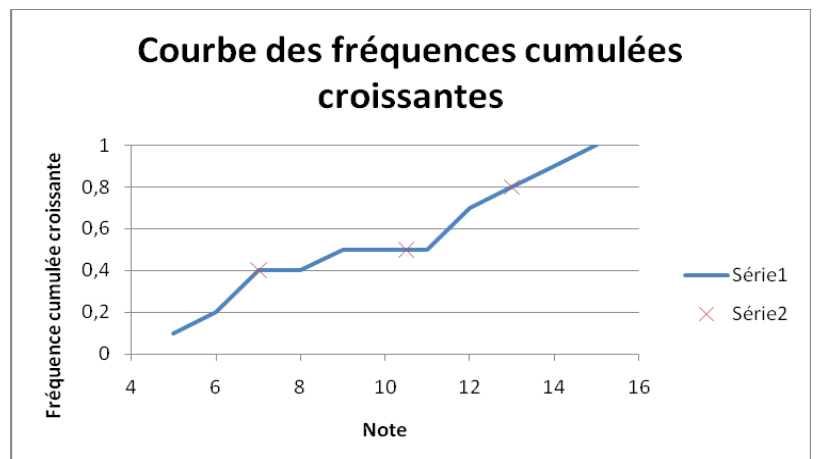
Me=11,5 $Q_1 = 7$ et $Q_3 = 13,5$

Remarque :

L'écart interquartile mesure la dispersion des valeurs autour de la médiane ; plus l'écart est petit, plus les valeurs de la série appartenant à l'intervalle interquartile sont concentrées autour de la médiane.

Contrairement à l'**étendue** (notée e) qui mesure l'écart entre la plus grande et la plus petite valeur, l'écart interquartile élimine les valeurs extrêmes qui peuvent être douteuses, cependant il ne tient compte que de 50% de l'effectif...

On reprend le cas des notes de 10 élèves pour introduire notion de courbe de fréquences cumulées croissantes.



Exemple 2 :

Cas d'un caractère quantitatif continu :

Le tableau donne la répartition d'une classe de 28 élèves sur le temps de révision d'un contrôle de mathématiques :

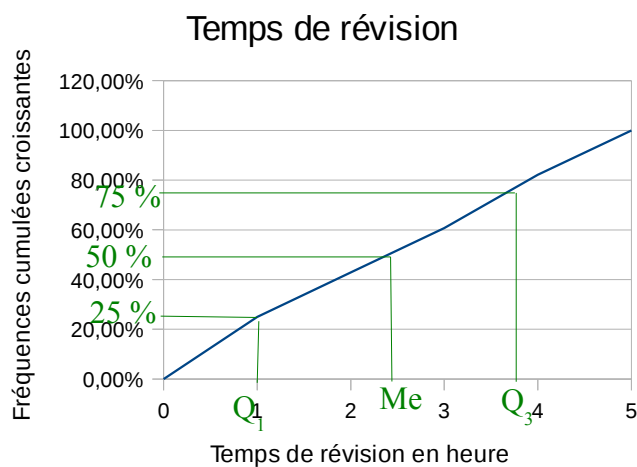
Temps de révision en h	$[0 ; 1[$	$[1 ; 2[$	$[2 ; 3[$	$[3 ; 4[$	$[4 ; 5[$
Effectifs	7	5	5	6	5
Effectifs cumulés croissants	7	12	17	23	28
Fréquences en %	25	17,9	17,9	21,4	17,8
Fréquences cumulées croissantes en %	25	42,9	60,8	82,2	100

$Me \approx 2,4$; $Q_1 = 1$ et $Q_3 \approx 3,7$

L'écart interquartile est $Q = Q_3 - Q_1 = 2,7$

Remarque :

On peut lire graphiquement sur la courbe des FCC, les quartiles et la médiane comme indiqué sur le graphique suivant.



De la même façon, on peut définir la courbe des fréquences cumulées décroissantes.