

LA STATISTIQUE

I- Généralité :

Population, caractère, valeur de caractère, effectif de la valeur, valeurs groupées par classes sont des expressions couramment utilisées en statistique.

Mais que signifie précisément chacun de ces termes ?

1- Population et individus :

Des élèves ont réalisé une **enquête statistique** dans leur collège. Ils ont demandé à leurs camarades de noter pendant une semaine le temps qu'ils avaient passé à regarder la télévision.

Dans l'enquête ci-dessus, les élèves du collège forment la **population** ; c'est sur eux que porte l'étude. Si on fait une étude statistique sur la couleur des voitures vendues en France pendant l'**année 1999**, la population sera l'ensemble de ces voitures.

Une population est formée d'**individus** ; dans l'enquête sur la télévision, un individu est un élève du collège. Dans le cas d'une étude sur la couleur de voitures, un individu est une voiture.

2- Caractères quantitatif et qualitatif :

Dans une étude statistique, le **caractère** est ce que l'on observe sur chaque individu. Pour l'enquête sur la télévision, le caractère est le **nombre d'heures** passées à regarder la télévision.

Ce caractère étant une **quantité**, on dit qu'il est **quantitatif**.

Dans le cas d'une étude sur la couleur des voitures, le caractère est la **couleur**. Ce caractère ne s'exprimant pas à l'aide d'un nombre, on dit qu'il est **qualitatif**.

3- Valeur :

L'étude d'un caractère sur une population fournit différents résultats (des réponses ou des observations) ; ces résultats sont les **valeurs du caractère**.

Dans l'enquête sur la télévision, les valeurs sont les nombres d'heures qui apparaissent sur la première ligne du tableau (0 h ; 1 h ; 2 h ; etc.).

Dans le cas d'une étude sur la couleur des voitures, les valeurs seront, par exemple, **blanc, gris, noir, vert, rouge, bleu**, etc.

4- Effectif :

L'effectif total est le nombre d'individus de la population. Dans l'enquête sur la télévision, l'effectif total est le nombre d'élèves interrogés ; on le trouve en additionnant les nombres de la deuxième ligne du tableau :

$$15 + 3 + 7 + 19 + 27 + 26 + 23 + 15 + 17 + 8 + 14 + 5 + 3 + 1 + 2 + 1 + 1 = 187$$

Dans le tableau, on voit que : 15 élèves n'ont pas regardé la télévision (0 heure), 3 élèves ont regardé la télévision 1 heure, 7 élèves ont regardé la télévision 2 heures, etc.

On peut traduire cela en disant que : l'effectif de la valeur 0 est égal à 15, l'effectif de la valeur 1 est égal à 3, l'effectif de la valeur 2 est 7, etc.

Remarque : il ne faut pas confondre l'effectif d'une valeur avec l'effectif total.

II- Résultats de l'enquête :

Les valeurs sont groupées par classes. On s'intéresse à présent à trois catégories d'élèves :

- ceux qui regardent la télévision moins de 7 heures par semaine (de 0 h à 6 h) ;
- ceux qui regardent la télévision au moins 7 heures mais moins de 14 heures (de 7 h à 13 h) ;
- ceux qui regardent la télévision au moins 14 heures par semaine (de 14 h à 16 h).

L'effectif 120, par exemple, a été obtenu en additionnant les effectifs des valeurs 0, 1, 2, 3, 4, 5 et 6 :
 $15 + 3 + 7 + 19 + 27 + 26 + 23 = 120$.

De même : $15 + 17 + 8 + 14 + 5 + 3 + 1 = 63$ et $2 + 1 + 1 = 4$.

Les trois regroupements de temps qui ont été faits s'appellent des **classes** :

- la 1^{re} classe est « de 0 h à 6 h » ; elle contient les valeurs 0, 1, 2, 3, 4, 5 et 6 ;

- la 2^e classe est « de 7 h à 13 h » ; elle contient les valeurs 7, 8, 9, 10, 11, 12 et 13 ;
- la 3^e classe est « de 14 h à 16 h » ; elle contient les valeurs 14, 15 et 16.

Remarque :

Ici, le mot classe ne doit pas être confondu avec celui qui désigne un groupe d'élèves (comme dans la phrase : « Dans cette classe de cinquième, les élèves ont un bon niveau »).

III- Représentation des données :

Dans une classe de 30 élèves, le professeur de mathématiques a établi le tableau ci-dessous qui donne les notes de ses élèves à un contrôle (le professeur note de 0 à 20 et toutes ses notes sont entières).

1- La Moyenne :

« J'ai la moyenne en anglais. » « Le professeur a dit que j'avais 12,5 de moyenne. »
Dans ces deux phrases, le mot *moyenne* n'a pas le même sens.

Dans la seconde phrase, 12,5 est le résultat d'un calcul, sans doute la somme de toutes les notes obtenues en anglais divisée par le nombre de notes, alors qu'« avoir la moyenne » signifie que le résultat du calcul précédent est supérieur à 10.

Il existe plusieurs manières de calculer une moyenne.

Quelles sont-elles ?

Si on souhaite une représentation plus globale des notes, on peut, **par exemple**, les partager en **trois classes** : de 0 à 7, de 8 à 12 et de 13 à 20.

2- Méthode de calcul :

Pour calculer la moyenne pondérée d'une série statistique :
on multiplie chaque valeur par l'effectif correspondant ;
on additionne tous les produits obtenus ;
on divise cette somme par l'effectif total.

IV- La Médiane :

Un élève qui a obtenu **7 sur 20** en maths cherche à relativiser l'insatisfaction de son père en lui faisant remarquer qu'il se situe exactement « en milieu de classe », c'est-à-dire qu'il y a autant d'élèves « devant lui » que « derrière lui ».

En statistique, on dit que la note de cet élève est une note médiane : comme en **géométrie**, le mot **médiane** est donc lié à la notion de milieu.

La médiane d'une série statistique est ce qu'on appelle une **caractéristique de position**.

Comment la déterminer ?

1- Définition :

Étant donnée une série statistique ordonnée, une valeur médiane de cette série est une valeur qui partage cette série en **deux groupes d'effectifs égaux** :

un groupe constitué de valeurs inférieures ou égales à la médiane ;

un groupe constitué de valeurs supérieures ou égales à la médiane.

Exemples :

- Cas où l'effectif est impair :

On veut déterminer une médiane de chacune des séries de nombres suivantes.

Première série : **2 ; 6 ; 7 ; 25 ; 58**.

Le nombre **7** est la valeur médiane de cette série : on a partagé la série en deux groupes d'effectifs égaux : **2 ; 6** (valeurs inférieures à **7**) et **25 ; 58** (valeurs supérieures à **7**).

Deuxième série : 4 ; 7 ; 9 ; 9 ; 11 ; 15 ; 17.

Le nombre **9** est la valeur médiane de cette série ; on a partagé la série en **deux groupes** d'effectifs égaux : **4 ; 7 ; 9** (valeurs inférieures ou égales à **9**) et **11 ; 15 ; 17** (valeurs supérieures à **9**).

Conclusion :

Si l'effectif est impair, la médiane est **unique** : c'est une valeur de la série.

Les **deux groupes** d'effectifs égaux sont alors constitués, d'une part, des valeurs qui précèdent la médiane dans la série ordonnée, d'autre part, des valeurs qui suivent la médiane dans la série ordonnée.

- Cas où l'effectif est pair :

On veut déterminer une médiane de chacune des séries de nombres suivantes.

Première série : 1 ; 5 ; 12 ; 13 ; 21 ; 24.

Le nombre 12,5 est **une** valeur médiane de cette série : on a partagé la série en **deux groupes** d'effectifs égaux : 1 ; 5 ; 12 (valeurs inférieures à 12,5) et 13 ; 21 ; 24 (valeurs supérieures à 12,5).

En fait, pour cette série, tout nombre strictement compris entre 12 et 13 est une valeur médiane.

Deuxième série : 5 ; 14 ; 18 ; 19 ; 19 ; 25 ; 47 ; 56.

Le nombre 19 est la valeur médiane de cette série ; on a partagé la série en **deux groupes** d'effectifs égaux : 5 ; 14 ; 18 ; 19 (valeurs inférieures ou égales à 19) et 19 ; 25 ; 47 ; 56 (valeurs supérieures à 19).

2- Détermination d'une médiane :

Il faut penser à ordonner la liste des valeurs de la série si celle-ci n'est pas ordonnée.

Exemple : voici la série ordonnée des notes (sur 20) obtenues par des élèves à un contrôle :

2 ; 3 ; 5 ; 6 ; 6 ; 8 ; 9 ; 9,5 ; 10 ; 10 ; 10 ; 11 ; 11 ; 12 ; 14 ; 14 ; 15,5 ; 16 ; 17 ; 17,5 ; 19.

En comptant les notes, on trouve un effectif total de 21. On va partager la série en **deux groupes** de 10 ; la note restante sera alors la note médiane de la série : il s'agit ici de la onzième note de la série, qui est égale à 10.

Conclusions :

Si l'effectif est pair, les **deux groupes** d'effectifs égaux correspondent aux **deux** moitiés de la liste des valeurs de la série ordonnée ; l'**exemple** de la **première** série montre que si l'effectif est pair, la médiane n'est pas nécessairement unique et n'est pas toujours une valeur de la série.