

Chapitre 8 : Statistiques et probabilités.

I. Révisions : pourcentages.

Ecriture en %	Ecriture fractionnaire	Ecriture décimale
18%	$\frac{18}{100}$	0,18
12,4%	$\frac{12,4}{100}$	0,124
0,61%	$\frac{0,61}{100}$	0,0061
0,02%	$\frac{0,02}{100}$	0,0002

En maths, le mot « de » signifie « \times ».

$$\begin{aligned} & 26\% \text{ de } 261\text{€} \\ &= 26\% \times 261\text{€} \\ &= 0,26 \times 261\text{€} \text{ (pourcentage en écriture décimale)} \\ &= 67,86\text{€} \end{aligned}$$

Remarque : en statistiques, un pourcentage s'appellera une fréquence.

II. Diagrammes

Exemple de série statistique :

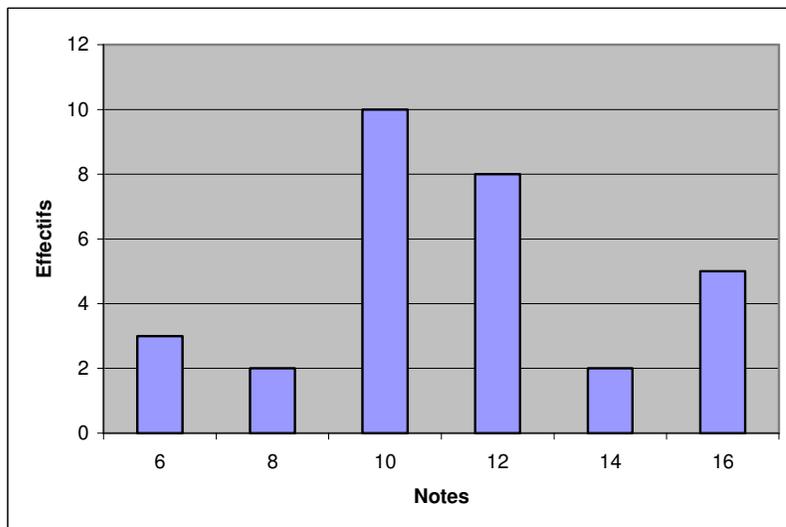
Notes (valeurs)	6	8	10	12	14	16	Effectif total :
Nombre d'élèves (effectif)	3	2	10	8	2	5	30

La série statistique est la liste de nombres suivante :

6 ; 6 ; 6 ; 8 ; 8 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 10 ; 12 ; 12 ; 12 ; 12 ; 12 ; 12 ; 12 ; 12 ; 14 ; 14 ; 16 ; 16 ; 16 ; 16 ; 16 .

A. Diagramme en bâtons ou histogramme :

La hauteur des « bâtons » est proportionnelle à l'effectif.



B. Diagramme circulaire

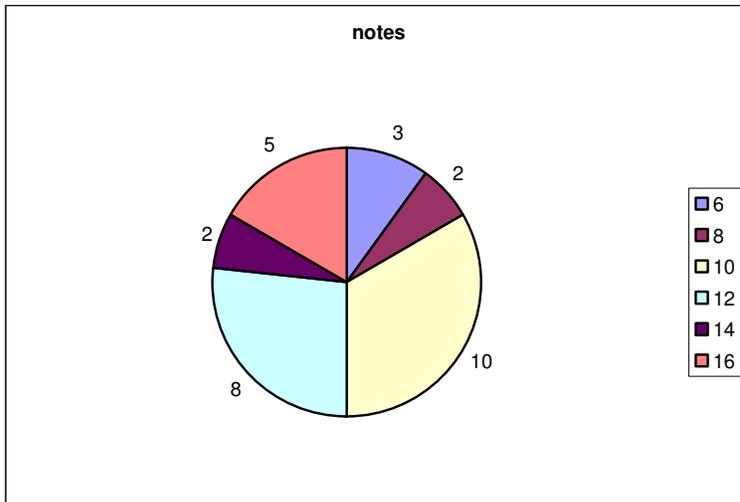
L'angle est proportionnel à l'effectif. L'effectif total représente 360°.

$$\text{Effectif total : } 3+2+10+8+2+5 = 30.$$

Effectif :	30	3	2	10	8	2	5) ×12
Angle :	360°	36°	24°	120°	96°	24°	60°	

Rappel : calcul du coefficient de proportionnalité « pour multiplier en descendant » :

Dans une colonne où je connais les deux nombres, coefficient = $\frac{\text{nombre du bas}}{\text{nombre du haut}} = \frac{360}{30} = 12$.



III. Effectifs cumulés croissants

Dans notre exemple :

Notes (valeurs)	6	8	10	12	14	16
Nombre d'élèves (effectif)	3	2	10	8	2	5
Effectif cumulé croissant	3	5	15	23	25	30

← On doit retrouver ici l'effectif total.

IV. Introduction au vocabulaire des statistiques : notations.

Une **série statistique** est une liste de nombres, par exemple la liste des notes obtenues par les élèves de 3^o au DS7 de mathématiques. Si deux élèves ont eu la même note, ce nombre apparaît deux fois dans la liste.

Les différentes notes obtenues s'appellent **les valeurs de la série statistiques**. S'il y a des élèves qui ont eu 12,5, on dit que 12,5 est une valeur de la série statistique.

Le **nombre** d'élèves qui ont eu 12,5 s'appelle **l'effectif** de la valeur 12,5.

Le **pourcentage** d'élèves qui ont eu 12,5 s'appelle **la fréquence** de la valeur 12,5.

Le nombre total d'élèves de 3^o s'appelle **l'effectif total**.

La « fréquence totale » est 100%.

V. Des effectifs aux fréquences, et inversement.

Pté 1 : Conversion d'un effectif en fréquence (on obtient un pourcentage en écriture décimale) :

$$\text{fréquence d'une valeur} = \frac{\text{effectif de la valeur}}{\text{effectif total}}$$

Pour avoir des fréquences en % : fréquence d'une valeur = $\frac{\text{effectif de la valeur}}{\text{effectif total}} \times 100$

Pté 2 : Conversion d'une fréquence (% en écriture décimale) en effectif :

$$\text{effectif de la valeur} = \text{fréquence de la valeur} \times \text{effectif total}$$

Avec des fréquences en % : effectif de la valeur = $\frac{\text{fréquence de la valeur} \times \text{effectif total}}{100}$

VI. Quelques grandeurs statistiques.

Ces grandeurs ne peuvent être calculées que si le caractère que l'on étudie est traduit par un nombre.

Exemple : Les joueurs de l'équipe de hand-ball mesurent en mètres:

1,58 - 1,65 - 1,65 - 1,68 - 1,75 - 1,78 - 1,80 - 1,90.

Ceci est une série statistique.

Def 1 : La **moyenne** d'une série statistique est la somme des valeurs (chaque valeur est comptée autant de fois que son effectif) divisée par l'effectif total.

$$moy = \frac{1,58 + 1,65 + 1,65 + 1,68 + 1,75 + 1,78 + 1,80 + 1,90}{8} \approx 1,72m$$

Def 2 : Après avoir rangé les nombres d'une série statistique dans l'ordre croissant, on appelle **médiane** de cette série un nombre qui partage la série en deux ensembles de même effectif : les nombres qui précèdent la médiane et les nombres qui suivent la médiane.

On peut prendre comme médiane de cette série n'importe quel nombre compris entre 1,68 et 1,75 (la médiane n'est pas forcément une valeur de la série statistique).

Par exemple une médiane de cette série est 1,70m : il y a autant de joueurs plus grands que 1,70m que de joueurs plus petits.

Remarques : contrairement à la moyenne, la médiane ne dépend pas des valeurs extrêmes (la plus grande et la plus petite). Cela peut être intéressant par exemple dans une expérience de physique, où les valeurs extrêmes sont souvent dues à des erreurs de manipulation.

La médiane et la moyenne sont deux caractéristiques de position de la série statistique.

Def 3 : Après avoir rangé les nombres d'une série statistique dans l'ordre croissant, on calcule l'**étendue** de cette série en effectuant la soustraction : (plus grande valeur) - (plus petite valeur).

L'étendue de cette série est :

$$\text{étendue} = 1,90 - 1,58 = 0,32$$

Remarque : L'étendue est une caractéristique de dispersion de la série statistique.

Def 4 : Soit une série statistique.

On appelle **premier quartile** la plus petite valeur de la série, notée Q_1 , telle qu'au moins 25% des valeurs de la série soient inférieures ou égales à Q_1 .

On appelle **troisième quartile** la plus grande valeur de la série, notée Q_3 , telle qu'au moins 75% des valeurs de la série soient inférieures ou égales à Q_3 .

La différence $Q_3 - Q_1$ s'appelle **écart interquartile**.

Le premier quartile de cette série est $Q_1 = 1,65$; le troisième quartile est $Q_3 = 1,78$.

L'écart interquartile est $1,78 - 1,65 = 0,13$.

Remarque : Le second quartile est la médiane : au moins 50% des valeurs de la série statistique sont inférieures ou égales à la médiane.

VII. Vocabulaire et premières notions de probabilités

Lorsqu'on fait une expérience, comme lancer un dé à 6 faces, dont on connaît tous les résultats possibles (les nombres de 1 à 6), mais que l'on ne sait pas à l'avance le résultat que l'on obtiendra, on dit que l'on fait une **expérience aléatoire**.

Chacun des résultats possibles s'appelle un **événement**.

La **probabilité** qu'un certain événement se produise représente les chances que cet événement se réalise lorsqu'on fait l'expérience aléatoire (chances d'obtenir un 6 lorsqu'on lance un dé), c'est un nombre compris entre 0 et 1, que l'on peut aussi voir comme un pourcentage en écriture décimale.

Si tous les événements possibles ont autant de chances d'arriver (**équiprobabilité**), la **probabilité** d'un événement se calcule ainsi : $\frac{\text{cas favorables}}{\text{cas possibles}}$.

Dans notre exemple, la probabilité d'obtenir 5 en tirant un dé est $p(5) = \frac{1}{6} \approx 0,166...$ soit environ 17% de chances.

La probabilité d'obtenir un 5 ou bien un 3 est $p(5 \text{ ou } 3) = p(5) + p(3) = \frac{2}{6} \approx 0,33...$ soit environ 33% de chances.

L'événement contraire de « obtenir 5 » est « ne pas obtenir 5 ». La probabilité de l'événement contraire se calcule ainsi : $p(\text{non } 5) = 1 - p(5) = 1 - \frac{1}{6} = \frac{5}{6} \approx 0,833...$

Lorsque l'expérience aléatoire comprend plusieurs événements successifs (lancer un dé, noter le résultat, puis le lancer une seconde fois), on dit que c'est une **expérience à deux épreuves**. La probabilité de l'événement « obtenir un 3 puis un 5 » est :

$$p(3 \text{ puis } 5) = p(3) \times p(5) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \approx 0,02778...$$

Si les deux événements sont liés, il faut en tenir compte :

$$p(\text{gagner au loto}) = p(\text{jouer au loto}) \times p(\text{gagner sachant qu'on joue}) .$$

On peut s'aider en faisant un « arbre des possibles » (voir exercices).