Chapitre 6

Statistiques

6.1 Exemple-modèle et vocabulaire de base

Voici un exemple qui va être utile pour poser les bases de vocabulaire de ce chapitre. On a demandé aux 29 élèves d'une classe de 3^{e} combien ils avaient de frères ou de surs. Les réponses obtenues sont :
1 ; 4 ; 0 ; 1 ; 1 ; 3 ; 3 ; 4 ; 2 ; 2 ; 1 ; 1 ; 3 ; 2 ; 2 ; 1 ; 0 ; 0 ; 2 ; 5 ; 2 ; 1 ; 1 ; 0 ; 1 ; 2 ; 1 ; 1 ; 0.

Définition 47. Chaque réponse d'élève est appelée une donnée.

Définition 48. La liste de ces 29 données est une série statistique.

Définition 49. La population étudiée est « les élèves de cette classe de 3^{e} ».

Définition 50. Le caractère étudié est « le nombre de frères ou de surs ».

Définition 51. Les valeurs prises par ce caractère sont les nombres 0 ; 1 ; 2 ; 3 ; 4 et 5.

Remarque. Il ne faut pas confondre les valeurs et les données. Il y a ici 29 données, mais seulement 6 valeurs.

6.2 Effectif et fréquence

Définition 52. L'effectif total d'une série statistique est le nombre de données dans la série.

Exemple 58. Dans l'exemple-modèle, l'effectif total est 29.

Définition 53. L'effectif d'une valeur donnée est le nombre de fois où cette valeur apparaît dans la série.

Exemple 59. Dans l'exemple-modèle, l'effectif de la valeur 2 est 7.

Remarque. Pour lire plus facilement les données d'une série statistique, on peut les regrouper et les organiser dans un tableau.

Nombre de frères et de surs	0	1	2	3	4	5
Effectif	5	11	7	3	2	1

La somme de tous les effectifs est égal à l'effectif total.

Définition 54. L'effectif cumulé croissant d'une valeur donnée est le nombre de fois où cette valeur ou une valeur inférieure apparaît dans cette série.

Exemple 60. Dans l'exemple-modèle, l'effectif cumulé croissant de la valeur 1 est 16.

Nombre de frères et de surs	0	1	2	3	4	5
Effectif	5	11	7	3	2	1
Effectif cumulé croissant	5	16				

L'effectif cumulé croissant de la valeur la plus élevée est l'effectif total.

Définition 55. La fréquence d'une valeur donnée est le quotient de l'effectif de la valeur par l'effectif total.
Une fréquence peut s'exprimer à l'aide d'une fraction, d'un pourcentage ou d'une écriture décimale.

Exemple 61. Dans l'exemple-modèle, l'effectif de la valeur 4 est 2, sur un effectif total de 29. La fréquence de la valeur 4 est donc $\dfrac{2}{29}$.

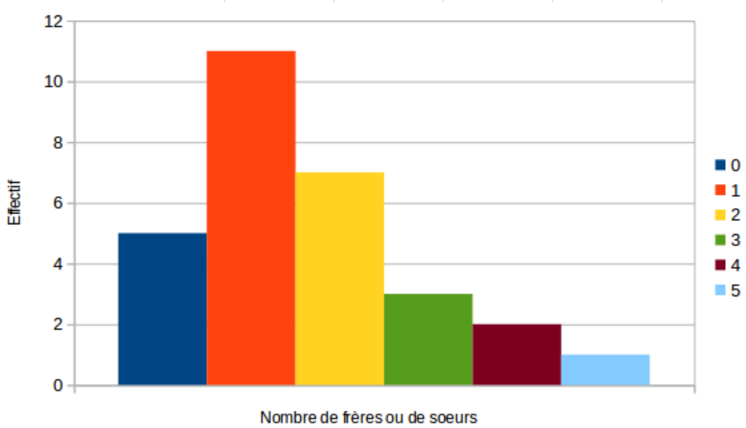
Nombre de frères et de surs	0	1	2	3	4	5
Effectif	5	11	7	3	2	1
Fréquence	$\dfrac{5}{29}$					

6.3 Représentation graphique

On peut représenter une série de données par un tableau, ou par différents types de graphiques : diagramme en barres ou en bâtons, circulaire (ou semi-circulaire), histogramme, courbe.

Propriété 56. *Dans un diagramme en bâtons, la hauteur de chaque bâton est proportionnelle à l'effectif de la valeur qu'il représente.*

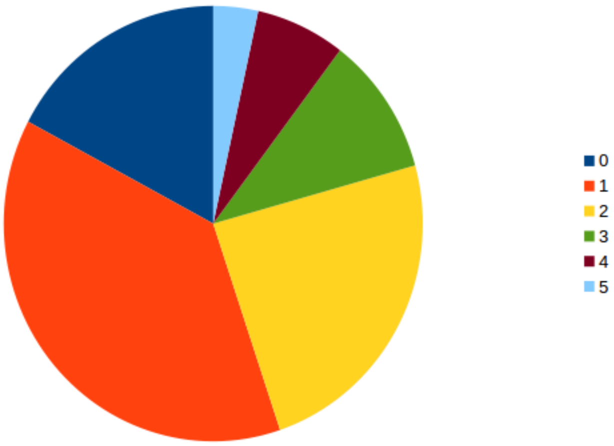
Remarque. Les bâtons sont espacés alors que les barres sont « collées ».



Propriété 57. *Dans un diagramme circulaire, la mesure de l'angle de chaque secteur est proportionnelle à l'effectif de la valeur qu'il représente.*

Exemple 62. Dans l'exemple-modèle, on calcule l'angle qui correspond à la valeur 4 : $\dfrac{2}{29} \times 360 \approx 25$

Nombre de frères et de surs	0	1	2	3	4	5
Effectif	5	11	7	3	2	1
Angle en degré	62					



Propriété 58. *Dans un histogramme, c'est l'aire de chaque rectangle qui est proportionnelle à l'effectif de la valeur qu'il représente.*

Remarque. Le terme « histogramme » est souvent employé à tort. Il s'agit le plus souvent de diagrammes en barres.

6.4 Moyenne et moyenne pondérée

Définition 56. La moyenne d'une série de données est le nombre égal à la somme des données divisée par l'effectif total de la série.

Exemple 63. Dans l'exemple-modèle, on a :
$$\frac{1+4+1+1+3+3+4+2+2+1+1+3+2+2+1+2+5+2+1+1+1+2+1+1}{29} = \frac{47}{29}$$
Chaque élève de la classe a donc, en moyenne, 1,62 frère ou sur.

On peut également utiliser une moyenne pondérée. On multiplie alors chaque valeur par son effectif : $\dfrac{1\times 11+2\times 7+3\times 3+4\times 2+5\times 1}{29} = \dfrac{47}{29}$

Exemple 64. Au bac, un élève a obtenu les notes suivantes : 9 à l'écrit de français (coefficient 5), 12 à l'oral de français (coefficient 5), 17 au grand oral (coefficient 10), 8 en philosophie (coefficient 8), 19 à son premier enseignement de spécialité (coefficient 16), et 15 à son deuxième enseignement de spécialité (coefficient 16). En classe de première, il avait 12 de moyenne dans son troisième enseignement de spécialité (coefficient 8), qu'il n'a pas poursuivi en terminale. Ses moyennes sur les deux années étaient de 11 en histoire-géographie (coefficient 6), 15 en EMC (coefficient 2), 13 en LV1 (coefficient 6), 15 en LV2 (coefficient 6), 12 en enseignement scientifique (coefficient 6) et 17 en EPS (coefficient 6).
A-t-il une mention ?

Calcul de sa moyenne :
$$\frac{(9+12)\times 5+17\times 10+8\times 8+(19+15)\times 16+12\times 8+11\times 6+15\times 2+(13+15+12+17)\times 6}{5+5+10+8+16+16+8+6+2+6+6+6+6} = \frac{1417}{100}$$
$$\frac{1417}{40} = 14,17.$$

L'élève obtient son baccalauréat avec mention bien.

6.5 Médiane et étendue

Définition 57. Une médiane d'une série de données est un nombre tel que :
• au moins la moitié des données sont inférieures ou égales à cette médiane ;
• au moins la moitié des données sont supérieures ou égales à cette médiane.

Remarque. • Pour trouver une médiane d'une série, il faut commencer par ranger les données dans l'ordre croissant.
• La médiane d'une série est une valeur de la série si l'effectif total est impair, mais pas nécessairement si l'effectif total est pair (on fait généralement la moyenne entre deux données).

Exemple 65. Si l'effectif total est impair, comme dans l'exemple-modèle :
effectif total = 29
$29 : 2 = 14,5$ donc la médiane est la 15^e donnée. On peut la trouver facilement grâce au tableau des effectifs cumulés croissants. La médiane de cette série est donc égale à 1. Cela signifie que plus de la moitié des élèves ont moins d'un frère ou sur et que la moitié de la classe a plus d'un frère ou sur.

Exemple 66. Déterminer une médiane de la série suivante : 5 ; 8 ; 9 ; 14 ; 18 ; 31.
L'effectif total est de 6, il est pair.
$6 : 2 = 3$ donc la médiane est située entre la 3^e et la 4^e donnée, donc entre 9 et 14. Une médiane de cette série est 11,5.

Exemple 67. Déterminer la médiane de la série suivante : 4 ; 7 ; 11 ; 12 ; 18 ; 26 ; 31.
L'effectif total est de 7, il est impair.
$7 : 2 = 3,5$ donc la médiane est la 4^e donnée, donc la médiane vaut 12.

Définition 58. L'étendue d'une série statistique est la différence entre les valeurs extrêmes de la série. Pour la calculer, on soustrait la plus petite valeur de la série à la plus grande valeur.

Exemple 68. Dans l'exemple précédent, l'étendue est égale à 26 car $31 - 5 = 26$.

Exemple 69. Dans l'exemple-modèle, l'étendue est égale à 5 car $5 - 0 = 5$.

Remarque. La médiane (comme la moyenne) est une caractéristique de position alors que l'étendue est une caractéristique de dispersion de la série.