2- La moyenne arithmétique

a) Définition

→ **La moyenne arithmétique simple**

La moyenne arithmétique d'une série ou **moyenne arithmétique simple** se calcule par une formule qui est donnée par l'expression :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Le "**x**" surmonté d'un trait désigne classiquement la moyenne. On remarque que la somme ci-dessus va de 1 à n où n désigne le nombre d'unités statistiques de la population.

Exemple :

Appliquons cette définition au calcul de la moyenne de la série suivante :

S= {4, 0, 1, 1, 2, 2, 2, 3, 3, 4, 2, 3, 4, 5, 2, 1, 3, 3, 4, 5}

On a donc :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{4+0+1+1+2+2+2+3+3+4+2+3+4+5+2+1+3+3+4+5}{20} = \frac{54}{20} = 2,7$$

Que signifie le chiffre 2,7 ?

2,7 est la valeur la plus proche à toutes les valeurs de la série. Autrement dit, il s'agit d'une valeur centrale, la plus proche à toutes les modalités de la série.

→ **La moyenne arithmétique pondérée**

La moyenne arithmétique d'une distribution ou **moyenne arithmétique pondérée** se calcule par une formule qui est donnée par l'expression :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$$

La somme varie cette fois de 1 à *k*, avec *k* qui représente le nombre de valeurs de la série. Dans le cas où aucune valeur n'est répétée $k=n$. Sinon $k<n$. Remarquons que la somme va de 1 à *k*, mais que cette somme est divisée par *n* et non par *k*.

La notation n_j représente les effectifs ou fréquences absolues des valeurs.

Exemple :

Appliquons cette définition au calcul de la moyenne de la distribution suivante :

x_j	n_j
0	1
1	3
2	5
3	5
4	4
5	2

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i = \frac{(0\times1)+(1\times3)+(2\times5)+(3\times5)+(4\times4)+(5\times2)}{20} = \frac{0+3+10+15+16+10}{20} = \frac{54}{20} = 2,7$$

b) Peut-on calculer la moyenne pour un caractère qualitatif ?

La moyenne arithmétique ne peut pas être calculée pour un caractère qualitatif. Soit par exemple le caractère sexe, avec les deux modalités "F" et "H", dans une population de 10 personnes. On a la série suivante : {F,H,F,F,H,H, F,F,F,H}. La modalité "Femme" est plus fréquente (6 contre 4 pour la modalité "Homme") : c'est le mode. En revanche, on ne peut pas calculer de moyenne arithmétique. La même chose est vraie si l'on met cette série sous forme d'une distribution :

Sexe	Effectif
H	4
F	6

On peut calculer les fréquences associées à chaque modalité. On voit alors que la modalité la plus fréquente est "F" (0,6 contre 0,4 pour la modalité "H"). Mais la notion de moyenne arithmétique n'a pas de sens.

c) La moyenne arithmétique pour un caractère quantitatif continu (classes)

Lorsque l'on veut calculer la moyenne d'une distribution par classes de valeurs, celle-ci s'obtient en prenant la formule de la moyenne pondérée et en remplaçant dans cette formule "x_j" par "c_j", où c_j représente le **centre de la classe j**, c'est-à-dire la moyenne arithmétique des extrémités de classe.

Exemple : soit la distribution statistique suivante :

Classes	n_j
[0-2[	4
[2- 4[	10
[4- 6]	6

Pour calculer la moyenne, nous devons déterminer les centres de classe, puis faire la somme des "n_j x c_j" et diviser par n. Autrement dit, nous devons appliquer la formule :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i$$

La notation **c_i** indique le centre de classe et où k représente le nombre de classes. Comme indiqué précédemment, le centre de classe est égal à la moyenne des extrémités de classe. On a donc le tableau de calcul suivant :

Classes	n_j	**c_j (moyenne des extrémités de classe)**	$n_j \times c_j$
[0-2[	4	1	4
[2- 4[	10	3	30
[4- 6]	6	5	30
			64

Et finalement :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i c_i = \frac{1}{20}\sum_{i=1}^{3}\left[(4\times1)+(10\times3)+(6\times5)\right] = \frac{4+30+30}{20} = \frac{64}{20} = 3,2$$