Statistique Descriptive

3ème Chapitre: Les caractéristiques de tendance centrale ou de position.

Introduction

L’objectif de ce chapitre est de présenter les principaux paramètres qui permettent de résumer une série statistique d’observations et d’éclairer sur la position du noyau (centre) de la série.

Ces paramètres permettent de savoir autour de quelles valeurs se situent les valeurs d’une variable statistiques. Ces paramètres sont appelés **caractéristiques de position** ou de **la tendance centrale** de la série statistique.

Nous présenterons ici le mode, la moyenne, la médiane et les quantiles.

Le mode

I. Le mode

Définition :

Le mode d'une variable statistique c'est la valeur la plus fréquente d'une série (la modalité qui représente le plus grand effectif).

Remarques :

* Le mode peut être calculé pour tous les types de variable : quantitative et qualitative.

* Le mode n'est pas nécessairement unique.

* Quand une variable continue est découpée en classes, on peut définir une classe modale (classe correspondant à l'effectif le plus élevé).

I. Le mode

1. Variable statistique qualitative ou quantitative discrète :

Exemple 1 : Si on considère la variable “Etat civil”, dont le tableau statistique suivant :

x_i	C	M	V	D
n_i	15	8	3	4

le mode est “C” : célibataire. L’effectif associé à ce mode : 15.

Exemple 2 :	**Exemple 3 :**
soit la série suivante :$\{9; 4; 6; 9; 3; 1\}$ La valeur la plus fréquente est : 9 Le mode est égal à 9 $\Downarrow$ Distribution unimodale	soit la série suivante : $\{9; 3; 6; 3; 5; 1; 5\}$ Les valeurs les plus fréquentes sont : 3 et 5 Cette série a 2 modes : 3 et 5. $\Downarrow$ Distribution bimodale

I. Le mode

2. Variable statistique quantitative continue :

Dans ce cas, l'effectif le plus élevé permet d'identifier la classe modale.

1. Classes d'amplitudes égales :

Exemple 4 :

Salaire	effectif	Amplitude
[2000,3000[	83	1000
[3000,4000[	102	1000
[4000,5000[	78	1000
[5000,6000[	117	1000
[6000,7000[	90	1000

- La classe modale c'est : [5000,6000[
- La classe inférieure à la classe modale : [4000,5000[
- La classe supérieure à la classe modale : [6000,7000[.

I. Le mode

Détermination du mode au sein de la classe modale

$$Mo = x_1 + \frac{k_1}{k_1+k_2}(x_2 - x_1)$$

Avec :

* $[x_1; x_2[$ c'est la classe modale.

*k_1 : la différence entre la fréquence (ou l'effectif) de la classe modale et de la classe précédente

*k_2 : la différence entre la fréquence (ou l'effectif) de de la classe modale et de la classe suivante.

*x_1 et x_2 représente respectivement la borne inférieure et supérieure de la classe modale.

I. Le mode

Comme application de cette formule, reprenons **l'exemple 4** précèdent :

- $[x_1; x_2[= [5000, 6000[$ c'est la classe modale.

$$k_1 = \text{l'effectif de la classe modale} - \text{l'effectif de la classe } [4000,5000[$$
$$k_1 = 117 - 78 = 39$$
$$k_2 = \text{l'effectif de la classe modale} - \text{l'effectif de la classe } [6000, 7000[$$
$$k_2 = 117 - 90 = 27$$

- L'amplitude $x_2 - x_1 = 1000$

Si on remplace ces quantités dans l'expression de "Mo", on trouve :

$$\text{Mo} = \mathbf{5000} + \frac{\mathbf{39}}{\mathbf{39} + \mathbf{27}} \times \mathbf{1000} = \mathbf{5590}$$

I. Le mode

2.2. Classes d'amplitudes inégales :

Pour des classes d'**amplitudes inégales**, la classe modale c'est la classe ayant l'effectif corrigé (ou la densité) le(a) plus élevé(e).

Exemple 5 : Reprenons un exemple vu dans le chapitre

précédent :

Classes	Effectif	Fréquence en %	Amplitude	densité en %
[6,9[	7	46,70	3	15,60
[9,11[	5	33,30	2	16,70
[11,14[	3	20,00	3	6,70

I. Le mode

Rappelons que la densité a pour expression $d_i = \frac{f_i}{a_i}$ et que

l'effectif corrigé = $d_i = \frac{n_i}{a_i}$

- Les amplitudes étant différentes, alors on utilise les densités pour déterminer la classe modale, ainsi $[x_1, x_2[= [9, 11[$.

k_1 = la densité de la classe modale − la densité de la classe [6,9[

$k_1 = 16.70 - 15.60 = 1.1$

k_2 = la densité de la classe modale − la densité de la classe [11,14[

$k_2 = 16.70 - 6.70 = 10$

- L'amplitude $x_2 - x_1 = 11 - 9 = 2$

Si on remplace ces quantités dans l'expression de "Mo", on trouve :

$$\text{Mo} = 9 + \frac{1.1}{1.1 + 10} \times 2 = 9.2$$

La moyenne

II. La moyenne

1.La moyenne arithmétique simple :

Définition :

La moyenne arithmétique simple d'une série statistique $(x_i)_{i=1}^{n}$ correspond à la division de leur somme par leur nombre. Elle se note $\bar{X}$, se calcule par la formule

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Exemple : Considérons la série $\{11, 12, 15, 12\}$, alors

$$\bar{x} = \frac{11 + 12 + 15 + 12}{4} = 12,5$$

II. La moyenne

2. La moyenne arithmétique pondérée :

Définition :

Lorsque les valeurs (modalités) sont affectées de coefficients (ici d'effectifs), on parle de moyenne arithmétique pondérée. Elle se calcule par la formule

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{r} n_i x_i = \sum_{i=1}^{r} f_i x_i ,$$

ou f_i est la fréquence de la modalité x_i et $\sum_{i=1}^{r} n_i = n.$

Exemple 1 : Voici les résultats d'un étudiant lors d'un concours

Matières	Maths	Économie	Français	anglais
coefficient	2	3	1	2
Notes	11.5	07	16	14

$$\bar{X} = \frac{2 \times 11.5 + 3 \times 7 + 1 \times 16 + 2 \times 14}{2 + 3 + 1 + 2} = 11$$

II. La moyenne

Exemple 2 :

Qualité se service	Effectif	$n_i x_i$	f_i	$f_i x_i$
1	1	1	0.08	0.08
2	3	6	0.25	0.5
3	5	15	0.42	1.26
4	2	8	0.17	0.68
5	1	5	0.08	0.4
Total (Somme)	12	**35**	1	**2.92**

la moyenne de la variable qualité de service $\overline{x} = \frac{35}{12} = \mathbf{2.92}$

II. La moyenne

La moyenne arithmétique pondérée sur variable continue en classe :

Pour calculer la moyenne arithmétique des données issues d'une variable continue et regroupées dans des classes, on applique la formule:

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

avec

- f_i est la fréquence de la la classe $[x_i, x_{i+1}[$.
- c_i est le centre de la classe $[x_i, x_{i+1}[$ donné par :

$$c_i = \frac{x_i + x_{i+1}}{2}$$

II. La moyenne

Exemple : On regroupe les notes des élèves de la manière suivante

Classes	n_i	f_i	c_i	$n_i \times c_i$	$f_i \times c_i$
[0 ; 5[	2	0.09	2.5	5	0.225
[5 ; 10[	3	0.14	7.5	22.5	1.05
[10 ; 12[	7	0.33	11	77	3.63
[12 ; 15[	5	0.24	13.5	67.5	3.24
[15 ; 17[	3	0.14	16	48	2.24
[17 ; 20]	1	0.05	18.5	18.5	0.925
Somme	21	1		238.5	**11.3**

La moyenne des notes c'est $\overline{x} = \frac{238.5}{21} = \mathbf{11.3}$

II. La moyenne

3.La moyenne géométrique simple :

Définition :

On définit la **moyenne géométrique** “G” de la façon suivante :

$$G = \sqrt[n]{x_1 \times x_2 \times \ldots \times x_n} = (x_1 \times x_2 \times \ldots \times x_n)^{\frac{1}{n}}$$

Remarque : L’une des applications de la moyenne géométrique c’est de trouver la moyenne de taux d’intérêt.

Exemple : Supposons que les taux d’intérêt pour 3 années consécutives soient respectivement de 5%; 10% et 15%. Que va-t-on obtenir après 3 ans si je empreinte 100 DHS ?

II. La moyenne

Réponse :

1- Après 1 an on a : $100 + 100 \times 5\% = 100 \times 1.05 = 105$ DHS ;

2- Après 2 ans on a :
$105 + 105 \times 10\% = 105 \times 1.1 = 100 \times 1.05 \times 1.1 = 115.5$DHS

3- Après 3 ans on a : $115.5 + 115.5 \times 15\% = 115.5 \times 1.15 =$
$100 \times 1.05 \times 1.1 \times 1.15 = 132.825$ DHS

Si on calcule la moyenne arithmétique des taux on obtient

$$\overline{x} = \frac{1.05 + 1.1 + 1.15}{3} = 1.1$$

Si on calcule la moyenne géométrique des taux, on obtient

$$G = (1.05 \times 1.1 \times 1.15)^{\frac{1}{3}} = 1.099241902$$

Le bon taux moyen est bien G et non $\overline{x}$ car si on applique 3 fois le taux moyen G aux 100 DHS, on obtient

$$100 \times G^3 = 100 \times (1.099241902)^3 = 132.825 \text{ DHS}$$

II. La moyenne géométrique:

Exemple 1:

1- Les bénéfices d’une entreprise ont augmenté:

- de 5% par an pendant les 2 premières années;
- de 9% par an pendant les 5 années suivantes;
- de 12% par an pendant les 3 années suivantes;

Quelle est l’augmentation moyenne sur les 10 ans?

Réponse 1 :

Soit B_0 le bénéfice réalisé par l'entreprise l'année précédant la période écoulée $B_0 \neq 0$. Le bénéfice était :

1 an plus tard : $B_1 = B_0 \times 1{,}05$

2 an plus tard : $B_2 = B_1 \times 1{,}05 = B_0 \times (1{,}05)^2$

3 ans plus tard : $B_3 = B_2 \times 1{,}09 = B_0 \times (1{,}05)^2 \times (1{,}09)$

...

10 ans plus tard : $B_{10} = B_0 \times (1{,}05)^2 \times (1{,}09)^5 \times (1{,}12)^3$

Soit a l'augmentation annuelle moyenne sur les 10 ans. On doit avoir :

$B_{10} = B_0 \times (1 + a)^{10}$

$(1 + a)^{10} = (1{,}05)^2 \times (1{,}09)^5 \times (1{,}12)^3$

$1 + a = [(1{,}05)^2 \times (1{,}09)^5 \times (1{,}12)^3]^{1/10}$ (moyenne géométrique pondérée des augmentations)

$\text{Ln}\,(1+a) = \frac{1}{10}\,[2\,\text{Ln}\,1{,}05 + 5\,\text{Ln}\,1{,}09 + 3\,\text{Ln}\,1{,}12]$

$\text{Ln}\,(1+a) \approx 0{,}0868 \Rightarrow 1+a) \approx e^{0{,}0868} \approx 1{,}09$.
Soit une augmentation de 9% par an, pendant 10 ans.

II. La moyenne

4.La moyenne harmonique : moyenne de l'inverse :

Définition :

On appelle **moyenne harmonique** la quantité:

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

Remarque : Il est judicieux d'appliquer la moyenne harmonique aux vitesses.

Exemple :

Un cycliste parcourt 4 étapes de 100km. Les vitesses respectives pour ces étapes sont de 10 km/h, 30 km/h, 40 km/h, 20 km/h.

Quelle a été sa vitesse moyenne ?

II. La moyenne

Réponse :

- Un raisonnement simple nous dit qu'il a parcouru la première étape en 10h, la deuxième en 3h20 la troisième en 2h30 et la quatrième en 5h. Il a donc parcouru le total des 400km en

10h + 3h20 + 2h30 + 5h = 20h50 = 20.8333h

sa vitesse moyenne est donc : Moy = $\frac{400}{20.8333}$ = 19.2Km/h.

- Si on calcule la moyenne arithmétique des vitesses, on obtient

$$\bar{x} = \frac{10 + 30 + 40 + 20}{4} = 25\text{Km/h}$$

- Si on calcule la moyenne harmonique des vitesses, on obtient

$$H = \frac{4}{\frac{1}{10} + \frac{1}{30} + \frac{1}{40} + \frac{1}{20}} = 19.2\text{Km/h}$$

Merci