Institut Supérieur des Etudes

Technologiques de Sfax

<u>Cours de</u> : Statistiques & Probabilités

<u>Chapitre IV</u> : Les caractéristiques de forme

<u>Enseignant</u> : Mr. Anis HDIDER

<u>Niveau</u> : 1ère Année

<u>Département</u> : Informatique

<u>Année universitaire</u> : 2023/2024

Plan du chapitre

- **Section I : La mesure de l'asymétrie**

  I – Définition

  II – Les coefficients d'asymétrie

  1- Le coefficient de Yule

  2- Le coefficient de Pearson

  3- Le coefficient de Fisher

- **Section II : La mesure de l'aplatissement**

  I – Définition

  II – Le coefficient d'aplatissement de Fisher

- **Série des exercices N° 4**

Les caractéristiques de forme

Introduction :

Les caractéristiques de forme décrivent l'apparence de la distribution d'une série de données sans nécessiter de tracer la courbe de fréquence. Elles fournissent des indications sur la régularité de la courbe et sur son aplatissement.

L'asymétrie renseigne sur la régularité de la courbe, indiquant si elle est symétrique ou non par rapport à sa moyenne. En revanche, l'aplatissement permet de mesurer à quel point la distribution est concentrée autour de sa moyenne. Ainsi, ces mesures fournissent des informations précieuses sur la forme de la distribution des données, sans avoir besoin de visualiser la courbe de fréquence.

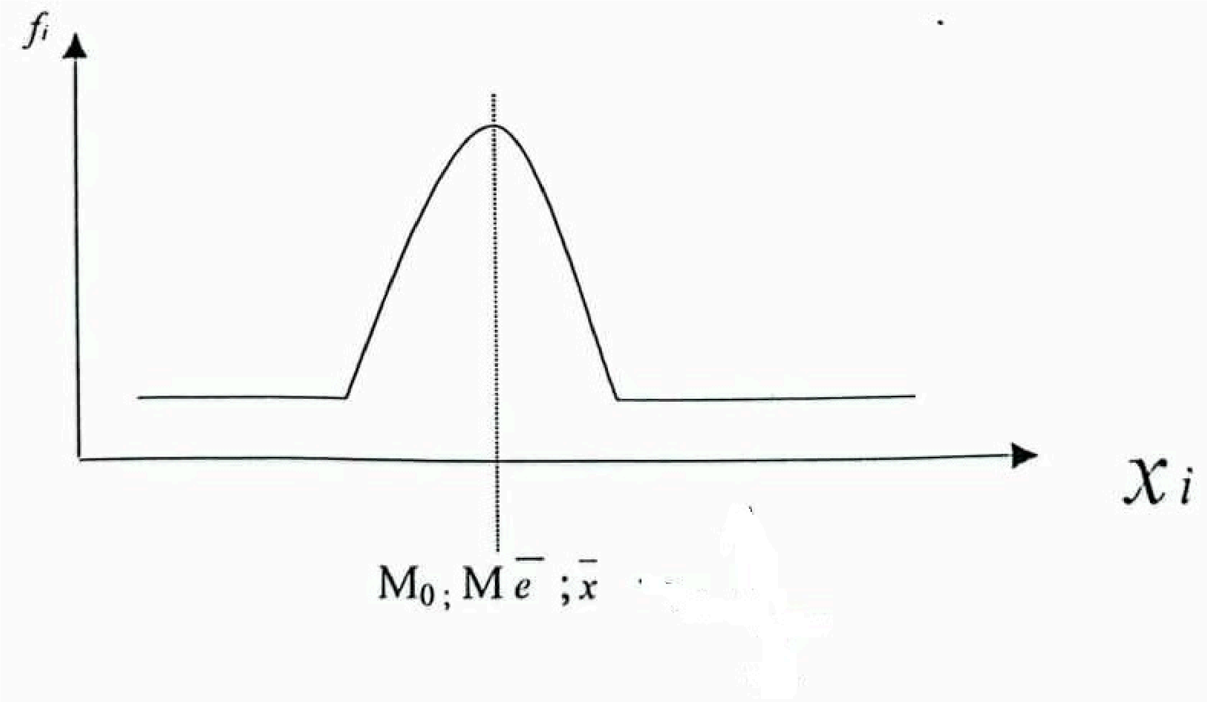
Section I : La mesure de l'asymétrie :

I- Définition :

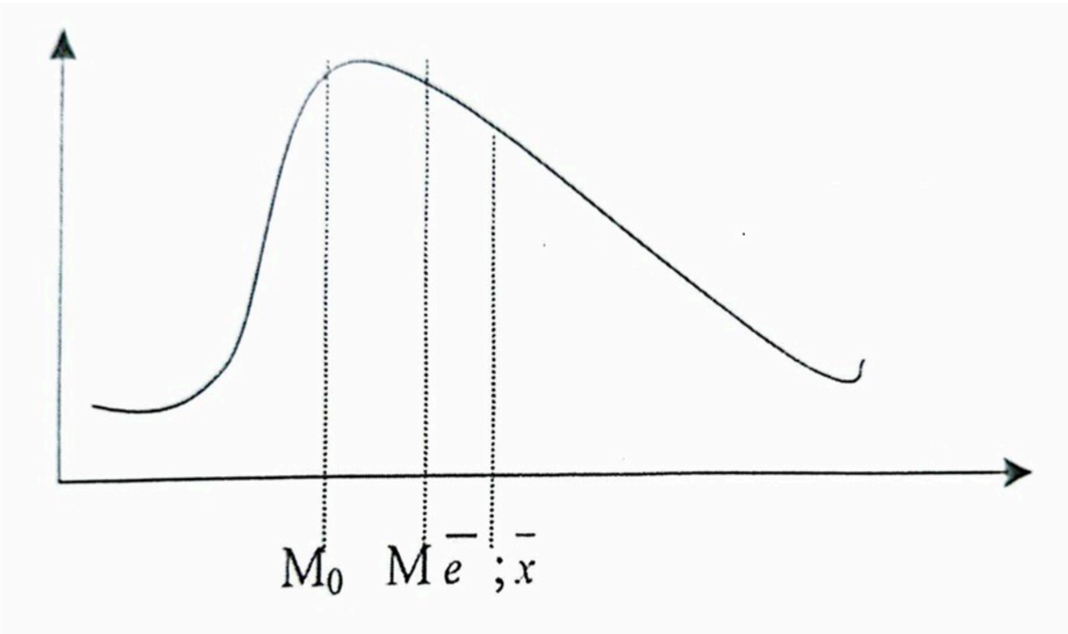
Une distribution est considérée comme symétrique lorsque les valeurs sont réparties uniformément de part et d'autre de la moyenne, de sorte que les valeurs extrêmes à gauche et à droite de la moyenne sont équidistantes. En revanche, une asymétrie positive indique que la queue de la distribution est plus étendue du côté des valeurs supérieures à la moyenne, tandis qu'une asymétrie négative indique que la queue de la distribution est plus étendue du côté des valeurs inférieures à la moyenne.

Dans une distribution symétrique, le mode, la médiane et la moyenne arithmétique sont confondues.

L'asymétrie peut être mesurée à l'aide de différents indices, tels que le coefficient de Yule, de Pearson, de Fisher ou de Skewness. Ces mesures fournissent des informations sur la forme et la distribution des données, permettant ainsi de mieux comprendre le comportement de la variable étudiée.



La courbe non symétrique est dite oblique.



II- Les coefficients d'asymétrie :

Dans ce chapitre on va s'intéresser à 3 coefficients pour la mesure de l'asymétrie, ces coefficients sont connus par les noms de leurs auteurs : **Yule, Pearson et Fisher**.

1- Le coefficient d'asymétrie de Yule :

C'est une mesure de l'asymétrie en comparant l'étalement vers la droite et l'étalement vers la gauche. Cette étalement est repérée par la position des quartiles (Q1, Mé, Q3). Le coefficient d'asymétrie de Yule s'écrit :

$$\boxed{C_Y = \frac{Q_3 + Q1 + 2Mé}{Q_3 - Q1}}$$

La valeur du coefficient de Yule est toujours comprise entre -1 et +1 et son signe indique le sens de l'asymétrie :

- Si $C_Y < 0$, la distribution est asymétrique (étalée à gauche ou oblique à droite).
- Si $C_Y = 0$, la distribution est symétrique.
- Si $C_Y > 0$, la distribution est asymétrique (étalée à gauche ou oblique à droite).

2- Les coefficients d'asymétrie de Pearson :

Karl Pearson a proposé deux coefficients de dissymétrie basés sur les écarts entre les mesures de tendance centrale.

- ✓ Le premier coefficient correspond à la différence entre la moyenne arithmétique et le mode divisé par l'écart type :

$$\boxed{\delta = \frac{\bar{X} - Mo}{\sigma(x)}}$$

- ✓ Le deuxième coefficient correspond à la différence entre la moyenne arithmétique et la médiane multipliée par 3, le tout est divisé par l'écart type de la distribution.

$$\boxed{\delta = \frac{3(\bar{X} - Mé)}{\sigma(x)}}$$

Les coefficients de Pearson s'interprètent comme le coefficient de Yule, c'est-à-dire :

➡ Si $\delta < 0$, la distribution est asymétrique (étalée à gauche ou oblique à droite).

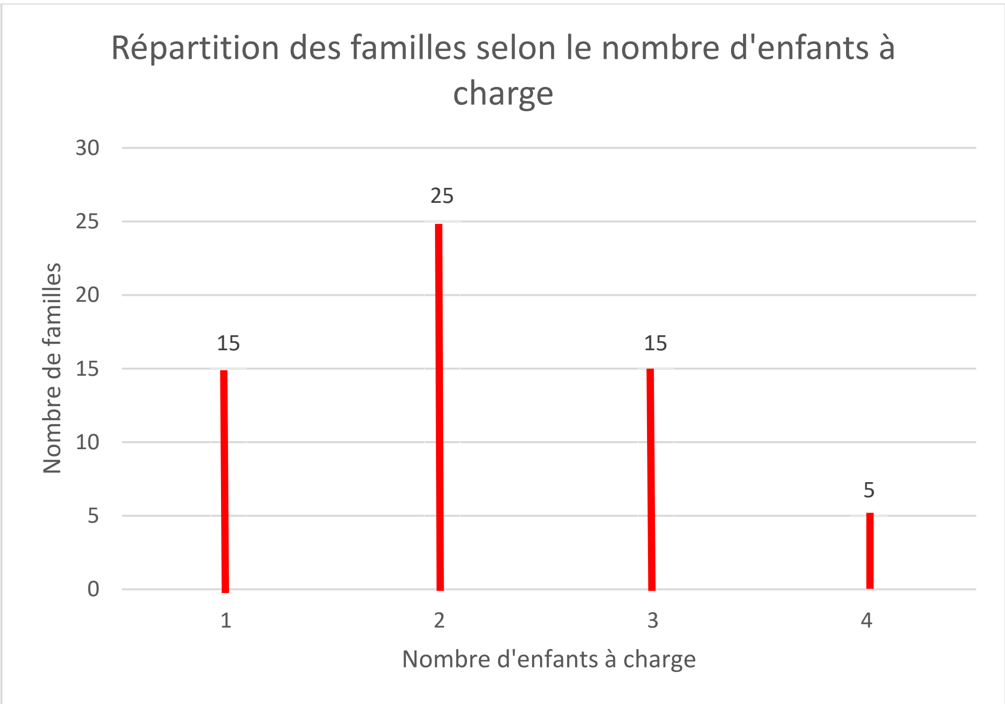
➡ Si $\delta = 0$, la distribution est symétrique.

➡ Si $\delta > 0$, la distribution est asymétrique (étalée à droite ou oblique à gauche).

Exemple :

Soit la répartition suivante de 60 familles selon le nombre d'enfants à charge :

Nombre d'enfants à charge	Effectif (ni)
1	15
2	25
3	15
4	5
Total	**60**



Graphiquement, on peut remarquer que la distribution est oblique à gauche, étalée à droite.

Le mode de la distribution est Mo = 2.

La moyenne arithmétique est : $\overline{X} = 2{,}167$.

Nombre d'enfants à charge	Effectif (ni)	Fréquence (fi)	N croissant	F croissant
1	15	0,25	15	0,25
2	25	0,417	40	0,667
3	15	0,25	55	0,917
4	5	0,083	60	1
Total	**60**	**1**		

- D'après le tableau statistique N(Q1)= 1/4*60 = 15, se trouve entre la première et la deuxième ligne, donc Q1 = 2.
- D'après le tableau statistique N(Q2)=N(Mé) = ½*60 = 30, se trouve entre la première et la deuxième ligne, donc Q2 = Mé = 2.
- D'après le tableau statistique N(Q3) = 3/4*60 = 45 se trouve entre la deuxième ligne et la troisième ligne, donc Q3 = 3.

➢ **Calculons maintenant le coefficient de Yule :**

$$\boldsymbol{C_Y = \frac{Q_3 + Q1 - 2M\acute{e}}{Q_3 - Q1}}$$

$C_Y = \frac{3+2-4}{3-1} = \frac{1}{2} > 0$ donc la distribution est asymétrique et est oblique à gauche.

➢ **En utilisant le premier coefficient de Pearson :**

$$\boldsymbol{\delta = \frac{\overline{X} - Mo}{\sigma(x)}}$$

$V(x) = \frac{1}{N}\sum_{i=1}^{k} n_i\,(x_i - \bar{X})^2 =$

Nombre d'enfants à charge	Effectif (ni)	$(x_i - \bar{X})$	$(x_i - \bar{X})^2$	$ni(x_i - \bar{X})^2$
1	15	-1,167	1,361889	20,428335
2	25	-0,167	0,027889	0,697225
3	15	0,833	0,693889	10,408335
4	5	1,833	3,359889	16,799445
Total	**60**	**1**	**5,443556**	**48,33334**

$V(x) = 48{,}33334/60 = 0{,}805$.

$\sigma(x) = \sqrt{V(x)} = \sqrt{0{,}805} = 0{,}897$.

Coefficient d'asymétrie de Pearson : $\delta = \frac{\overline{X} - Mo}{\sigma(x)} = \frac{2{,}167 - 2}{0{,}897} = 0{,}186 > 0$

La distribution est asymétrique et est oblique à gauche.

- **En utilisant le deuxième coefficient de Pearson :** $\boxed{\delta = \frac{3(\overline{X} - Mé)}{\sigma(x)}}$

$\delta = \frac{3(\overline{X} - Mé)}{\sigma(x)} = \frac{3(2{,}167 - 2)}{0{,}897} = 0{,}558 > 0.$

La distribution est asymétrique et est oblique à droite.

3- Le coefficient de Fisher :

La formule proposé par Ronald Fisher pour la mesure de l'asymétrie est la suivante :

$$\boxed{F = \frac{\frac{1}{N}\sum_{i=1}^{k} n_i\,(x_i - \bar{X})3}{\sigma(x)\,3}}$$

→ Si F < 0, la distribution est oblique à droite (étalée à gauche).

→ Si F = 0, la distribution est symétrique.

→ Si F > 0, la distribution est oblique à gauche (étalée à droite).

Exemple : Revenons à l'exemple précédent pour calculer le coefficient de Fisher :

$$F = \frac{\frac{1}{N}\sum_{i=1}^{k} n_i\,(x_i - \bar{X})3}{\sigma(x)\,3}$$

Nombre d'enfants à charge	Effectif (ni)	$(x_i - \bar{X})$	$(x_i - \bar{X})^3$	$ni(x_i - \bar{X})^3$
1	15	-1,167	-1,589324463	-23,83986695
2	25	-0,167	-0,004657463	-0,116436575
3	15	0,833	0,578009537	8,670143055
4	5	1,833	6,158676537	30,79338269
Total	**60**	**1**	**5,142704148**	**15,50722222**

F = $\dfrac{\frac{1}{N}\sum_{i=1}^{k} n_i\,(xi - \bar{X})3}{\sigma(\mathrm{x})\ 3} = \dfrac{\frac{1}{60}\times 15{,}507}{0{,}722} = 0{,}358 > 0$, donc la distribution est oblique à gauche.

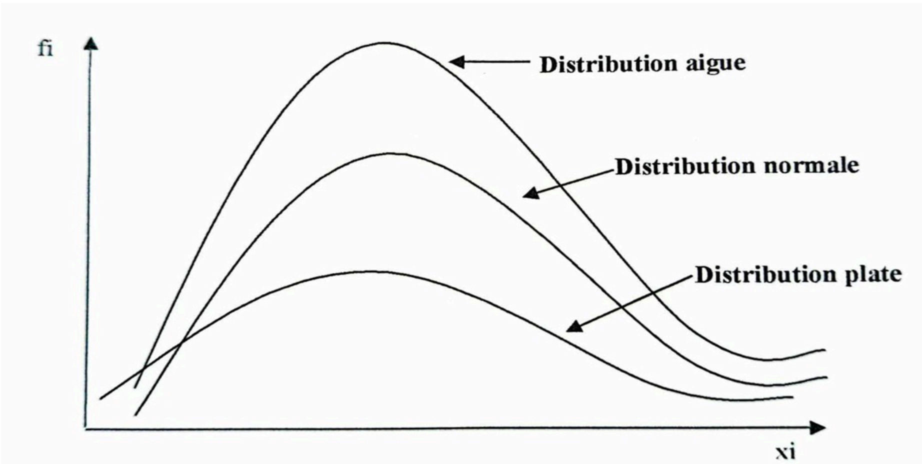
Section II : La mesure de l'aplatissement :

I- Définition :

L'aplatissement est une mesure de la forme de la distribution des données. Plus précisément, il mesure à quel point la distribution est « aplatie » par rapport à une distribution normale.

Une distribution normale a un aplatissement égal zéro. Si l'aplatissement est supérieur à zéro, la distribution est plus « pointue » et plus concentrée autour de la moyenne que la distribution normale, ce qui signifie qu'elle a des queues plus épaisses et des pics plus prononcés. Si l'aplatissement est inférieur à zéro, la distribution est plus plate que la distribution normale, avec des queues plus minces et des pics moins prononcés.

En résumé, l'aplatissement mesure à quel point les valeurs d'une distribution sont concentrées près de la moyenne et à quel point elles s'écartent des valeurs extrêmes.



II- Le coefficient d'aplatissement de Fisher :

Pour mesurer l'aplatissement, nous utilisons souvent le coefficient d'aplatissement de Fisher. Ce coefficient mesure à quel point la distribution est aplatie par rapport à une distribution normale.

Le coefficient d'aplatissement de Fisher est donné par la formule suivante :

$$\boxed{\beta = \frac{\frac{1}{N}\sum_{i=1}^{k} n_i\,(xi - \bar{X})4}{\sigma(x)\;4} - 3}$$

→ Si $\beta > 0$, la distribution est dite aigue ou leptocurtique.

→ Si $\beta = 0$, la distribution est dite normale ou nesocurtique.

→ Si $\beta < 0$, la distribution est dite plate ou platicurtique.

Exemple : Revenons à l'exemple précédent pour calculer le coefficient d'aplatissement de Fisher.

Nombre d'enfants à charge	Effectif (ni)	$(x_i - \bar{X})$	$(x_i - \bar{X})^4$	$ni(x_i - \bar{X})^4$
1	15	-1,167	1,854741648	27,82112472
2	25	-0,167	0,000777796	0,019444908
3	15	0,833	0,481481944	7,222229165
4	5	1,833	11,28885409	56,44427046
Total	**60**	**1**	**13,62585548**	**91,50706926**

$\beta = \frac{\frac{1}{N}\sum_{i=1}^{k} n_i\,(xi - \bar{X})4}{\sigma(x)\;4} - 3 = \frac{\frac{1}{60}\times 91{,}507}{0{,}647} - 3 = -0{,}643 < 0,$ la distribution est dite plate ou platicurtique.

Série d'exercices N° 4

Exercice N° 1 :

Soit la série suivante :

x_i	0	1	2	3
fi	0,216	0,432	0,288	0,064

1) Tracer la courbe des fréquence. En déduire graphiquement la forme de la distribution.
2) Déterminez les coefficients d'asymétrie et d'aplatissement de Pearson et de Fisher. Interprétez.

Exercice N° 2 :

Soit la série suivante :

Classe x_i	Effectif n_i	Centre ci	n × ci	N(x)
[10 ; 20 [	9	15	135	9
[20 ; 40 [	26	30	780	35
[40 ; 50 [	19	45	855	54
[50 ; 80 [	24	65	1560	78
[80 ; 100 [	14	90	1260	92
Total	**92**		**4590**	

1) Calculez le coefficient d'asymétrie de Yule.
2) Calculez le coefficient d'aplatissement de Fisher.