Université Moulay Ismail
Ecole Nationale de Commerce et de Gestion
-Meknès-

STATISTIQUE DESCRIPTIVE

-Semestre 2-

Enseignante : Noura AIT ELMEKKI

Etablissement : ENCG MEKNES

Année Universitaire : 2019-2020

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

(Suite 2)

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

- En plus des mesures de dispersion, de position et de concentration présentées dans les sections précédentes, on peut décrire une variable statistique en caractérisant la forme de sa distribution au moyen d'un indice.
- Cette forme est mise en évidence par une représentation graphique.
- Les indicateurs ou indices de forme donnent une idée de l'***asymétrie*** (en anglais, *skewness*) et de l'***aplatissement*** (en anglais, *kurtosis*) d'une distribution.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

- Pour caractériser la forme d'une distribution, c'est-à-dire pour préciser l'allure de la courbe des fréquences, il existe des coefficients permettant d'évaluer l'asymétrie d'une distribution et son aplatissement.
- ***Coefficients d'asymétrie***: *coefficient d'asymétrie de Fisher, coefficient d'asymétrie de Yule, et coefficient d'asymétrie de Pearson.*
- ***Coefficients d'aplatissements***: *coefficient d'aplatissement de Pearson et le coefficient d'aplatissement de Fisher.*

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie ou dissymétrie

- Une distribution statistique est symétrique si les observations repérées par leurs fréquences sont également dispersées de part et d'autre d'une valeur centrale.
- Dans une distribution symétrique, le graphe de la distribution (histogramme ou diagramme en bâton en fréquences) admet un axe de symétrie: le polygone des fréquences est symétrique par rapport à un axe vertical passant par son sommet.

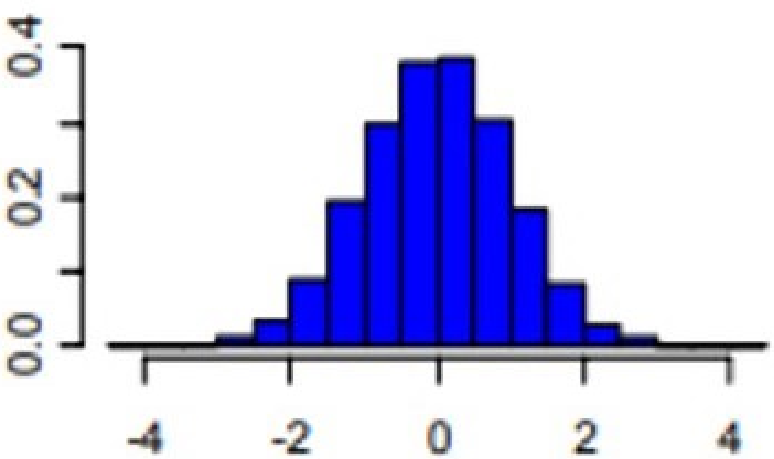
CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

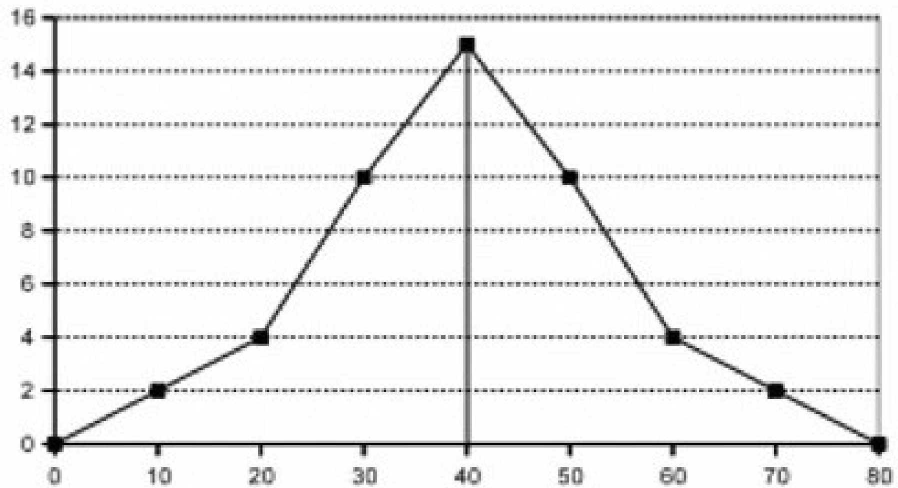
5. Caractéristiques de forme

L'asymétrie

- Lorsqu'une distribution est parfaitement symétrique, le mode, la médiane et la moyenne se confondent.

Distribution symétrique





$$M_e = \bar{x} = M_o$$

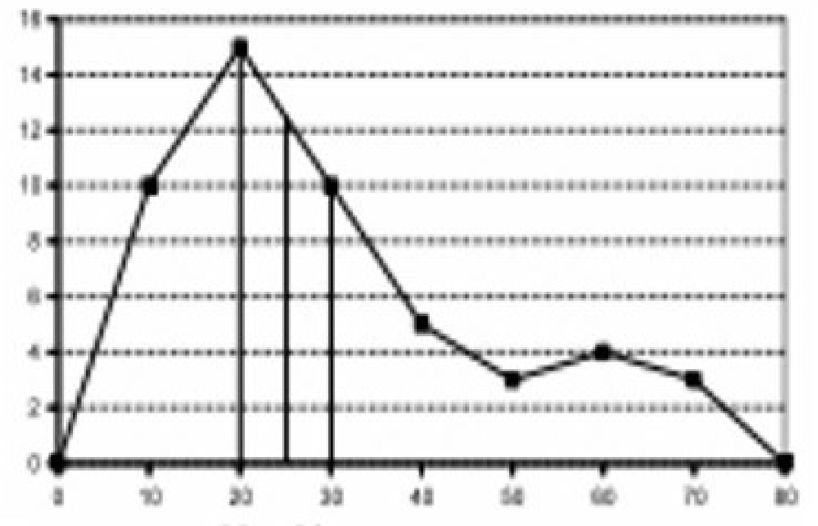
CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

- L'étalement des valeurs de la distribution peut être plus accentué à droite. Dans ce cas, la distribution des valeurs est ***asymétrique à droite***.

la portion du polygone des fréquences située à droite du sommet est plus longue que l'autre.



$$M_o < M_e < \bar{x}$$

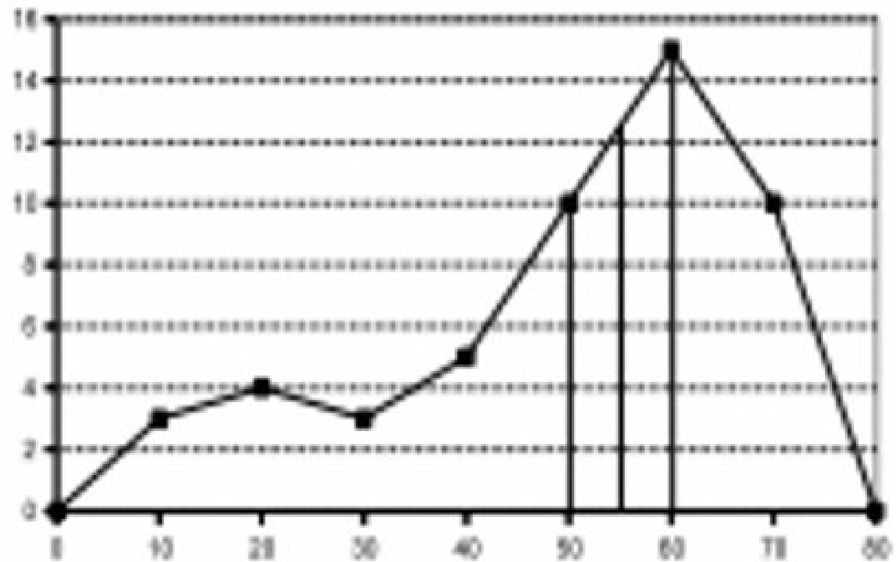
5. Caractéristiques de forme

L'asymétrie

- Au contraire, l'étalement des valeurs de la distribution peut être plus accentué à gauche. Dans ce cas, la distribution des valeurs est ***asymétrique à gauche***.

la portion du polygone des fréquences située à gauche du sommet est plus longue que l'autre.

$$M_o > M_e > \bar{x}$$

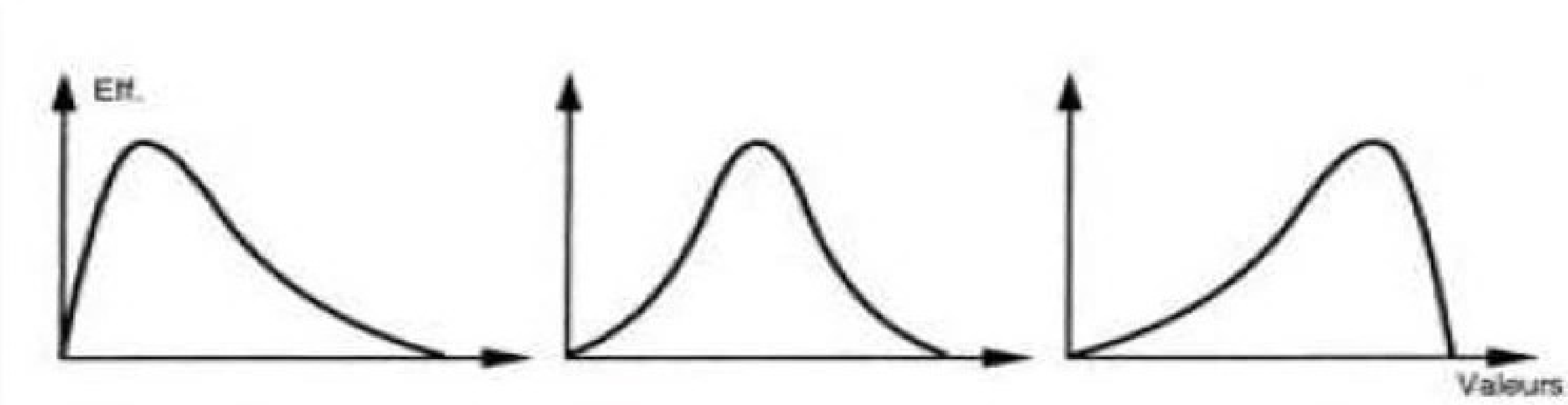


5. Caractéristiques de forme

L'asymétrie

- On distingue donc trois types de distributions selon qu'elles sont ***asymétriques à gauche*** (graphique de droite), ***symétriques*** (graphique du milieu) ou ***asymétriques à droite*** (graphique de gauche).

Asymétrie d'une distribution



$M_o > M_e > \bar{x}$ $\quad$ $M_e = \bar{x} = M_o$ $\quad$ $M_o < M_e < \bar{x}$

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Fisher

- Fisher a proposé un coefficient basé sur les écarts par rapport à la moyenne des valeurs en utilisant le moment centré d'ordre 3.

moment centré d'ordre 3

Indice d'asymétrie de Fisher $\longrightarrow$ $$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

cube de l'écart-type

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

Les moments

- Les moments sont des quantités algébriques qui permettent de décrire les caractéristiques des distributions statistiques : forme, symétrie, aplatissement, tendance centrale, dispersion.
- La moyenne arithmétique et la variance sont les moments les plus utilisés.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

Les moments

Moment d'ordre r simple (ou moment) non centré

- Le moment simple d'ordre *r* d'une série de *n* nombres x_1, x_2,. . . ,x_i,. . . ,x_n et de *k* effectifs n_1,n_2,….,n_k , est défini par :

$$m_r = \frac{1}{N}\sum_{i=1}^{k} n_i x_i^r$$

- Lorsque **r=1**, $\mathbf{m_1=\bar{x}}$: le moment non centré d'ordre 1 correspond donc à la moyenne arithmétique.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

Les moments

Moment centré d'ordre r

- Le moment centré d'ordre r d'une distribution est égal à la moyenne arithmétique des puissances d'ordre r des écarts $(x_i - \bar{x})$.

$$\mu_r = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^r = \sum_{i=1}^{k} f_i(x_i - \bar{x})^r$$

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

Les moments

Moment centré d'ordre r

- Le moment centré d'ordre 1 est nul.

$$\mu_1 = 0 \quad \text{car :} \quad \sum (x_i - \bar{x}) = \sum x_i - n\bar{x} = n\bar{x} - n\bar{x} = 0$$

- Le moment centré d'ordre 2 correspond à la variance.

$$\mu_2 = \frac{1}{N} \sum_{i=1}^{k} n_i (x_i - \bar{x})^2 = V(x)$$

$$\mu_2 = V(x) = \sigma_x^2$$

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Fisher

Indice d'asymétrie de Fisher

moment centré d'ordre 3

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\sum_{i=1}^{k} f_i (x_i - \bar{x})^3}{\sigma^3}$$

cube de l'écart-type

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Fisher

- Ce coefficient peut prendre des valeurs positives, négatives ou nulles:

  → Si $\gamma_1 < 0$, la distribution est asymétrique (étalée) à gauche.

  → Si $\gamma_1 = 0$, la distribution est symétrique.

  → Si $\gamma_1 > 0$, la distribution est asymétrique (étalée) à droite.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Yule

- Le coefficient de Yule sert à mesurer l'asymétrie de la distribution en tenant compte des positions relatives des quartiles par rapport à la médiane.
- Le coefficient d'asymétrie de Yule est basé sur les positions des 3 quartiles (1er quartile, médiane et troisième quartile), et est normalisé par la distance interquartile :

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Yule

$$C_Y = \frac{Q_3 + Q_1 + 2M_e}{Q_3 - Q_1}$$

- 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.

→Si $C_Y = 0$, la distribution est symétrique.

→Si $C_Y > 0$, la distribution est asymétrique (étalée) à droite.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Pearson

- Pearson a proposé de définir un coefficient de dissymétrie basé sur les écarts entre les mesures de tendance centrale.
- Il existe deux coefficients d'asymétrie dûs à Pearson.

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Pearson

- Le premier correspond à la différence entre moyenne et mode (ou moyenne et médiane), divisée par l'écart-type.

$$\delta = \frac{\bar{x} - M_O}{\sigma}$$

Comparaison entre les valeurs de la moyenne et du mode.

$$\delta = \frac{3(\bar{x} - M_e)}{\sigma}$$

Pearson a observé que dans les distributions des valeurs modérément asymétriques, la distance entre la moyenne et le mode est approximativement le triple de la distance entre la moyenne et la médiane.

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Pearson

- Il s'interprète comme le coefficient de Yule :

  →Si $\delta < 0$, la distribution est asymétrique (étalée) à gauche.

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

  →Si $\delta > 0$, la distribution est asymétrique (étalée) à droite.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Coefficient d'asymétrie de Pearson

- Le deuxième correspond au carré du coefficient de Fisher.
- Il est défini à partir des moments centrés d'ordre 2 et 3. C'est le moment d'ordre 3 au carré divisé par le moment d'ordre 2 au cube.

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

Où μ_2 n'est autre que la variance

- Le coefficient β_1 est toujours positif ou nul. Dans le cas où il est nul, il y a symétrie. Sinon, la distribution est oblique et tout dépend du signe de μ_3 : par exemple, si $\mu_3 > 0$, la distribution est asymétrique (ou étalée) à gauche.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Exemple

- *Une enquête a été mené auprès de 1500 ménages d'une région rurale par rapport à leurs nombre d'enfants. Le tableau suivant correspond à la distribution observée du nombre d'enfants des 1500 ménages de la région.*

Nombre d'enfants	Effectifs
1	380
2	455
3	245
4	230
5	100
6	75
7	10
8	5
Total	**1500**

Calculer les coefficients d'asymétrie de Fisher, Yule et Pearson

5. Caractéristiques de forme

L'asymétrie

Exemple : coefficient de Fisher

$$\gamma_1 = \frac{\mu_3}{\sigma^3} \qquad \text{où } \mu_3 = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^3$$

1. Calcul de μ_3

xi	ni	xi ni	$xi - \bar{x}$	$(xi - \bar{x})^3$	$ni(xi - \bar{x})^3$
1	380	380	-1.67	-4.65	-1767
2	455	910	-0.67	-0.30	-136.5
3	245	735	0.33	0.03	7.35
4	230	920	1.33	2.35	540.5
5	100	500	2.33	12.65	1265
6	75	450	3.33	36.92	2769
7	10	70	4.33	81.18	811.8
8	5	40	5.33	151.42	757.1
Total	**1500**	**4005**	-	-	**4247.25**

$\bar{x} = 4005/1500$

$\bar{x} = 2.67$

$\mu_3 = 4247.25/1500$

$\mu_3 = 2.83$

5. Caractéristiques de forme

L'asymétrie

Exemple : coefficient de Fisher $\gamma_1 = \frac{\mu_3}{\sigma^3}$ $\quad V(x) = \mu_2 = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^2$

2. Calcul de σ

xi	ni	xi ni	$xi - \bar{x}$	$(xi - \bar{x})^2$	$ni(xi - \bar{x})^2$
1	380	380	-1.67	2.79	1060.2
2	455	910	-0.67	0.45	204.75
3	245	735	0.33	0.11	26.95
4	230	920	1.33	1.77	407.1
5	100	500	2.33	5.43	543
6	75	450	3.33	11.09	831.75
7	10	70	4.33	18.75	187.5
8	5	40	5.33	28.41	142.05
Total	**1500**	**4005**	-	-	**3403.3**

$\bar{x}$ =4005/1500

$\bar{x}$ =*2.67*

μ_2=3403.3/1500

μ_2=*2.27*

$\sigma = \sqrt{2.27}$

σ =*1.50*

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Exemple : coefficient d'asymétrie de Fisher

3. Calcul du coefficient γ_1

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{2.83}{1.50^3} = \mathbf{0.83}$$

- $\gamma_1 > 0$, *la distribution du nombre d'enfants des 1500 ménages de la région est étalée à droite.*

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Exemple : coefficient d'asymétrie de Yule

$$C_Y = \frac{Q_3 + Q_1 + 2M_e}{Q_3 - Q_1}$$

Calcul du coefficient

xi	ni	ECC
1	380	380
2	455	835
3	245	1080
4	230	1310
5	100	1410
6	75	1485
7	10	1495
8	5	1500
Total	**1500**	-

$\mathbf{Q_1=1; M_e=2; Q_3=4}$

$\mathbf{C_Y=0.33}$

$C_Y>0$, la distribution du nombre d'enfants des 1500 ménages de la région est étalée à droite.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Exemple : coefficient d'asymétrie de Pearson $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$

Calcul du coefficient

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{2.83^2}{2.27^3} = \mathbf{0.68}$$

- *$\beta_1 > 0$, la distribution du nombre d'enfants des 1500 ménages de la région est étalée à droite.*

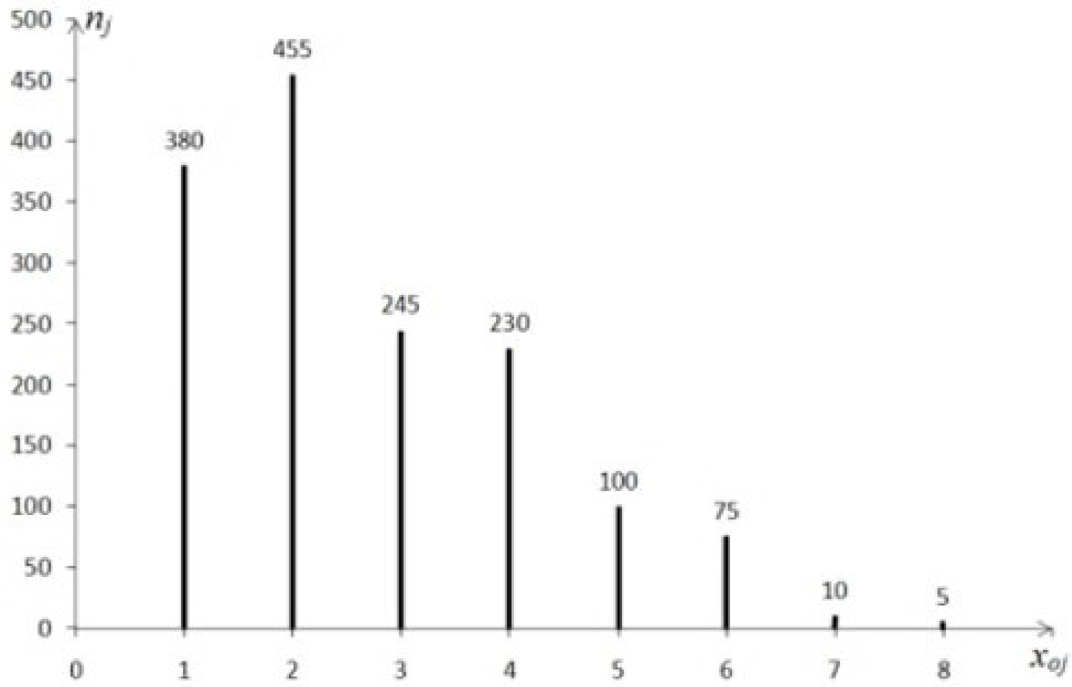
CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'asymétrie

Exemple

Nombre d'enfants	Effectifs
1	380
2	455
3	245
4	230
5	100
6	75
7	10
8	5
Total	**1500**



5. Caractéristiques de forme

L'aplatissement

- Les mesures d'aplatissement font partie des mesures qui caractérisent la forme d'une distribution.
- Kurtosis (du grec *kurtos* signifiant courbe ou arrondi) est une statistique descriptive mesurant l'aplatissement de la distribution ou ce qu'on appelle encore son degré de voussure.
- Les moments d'ordre 4 renseignent sur le degré d'aplatissement de la courbe de fréquences d'une distribution.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

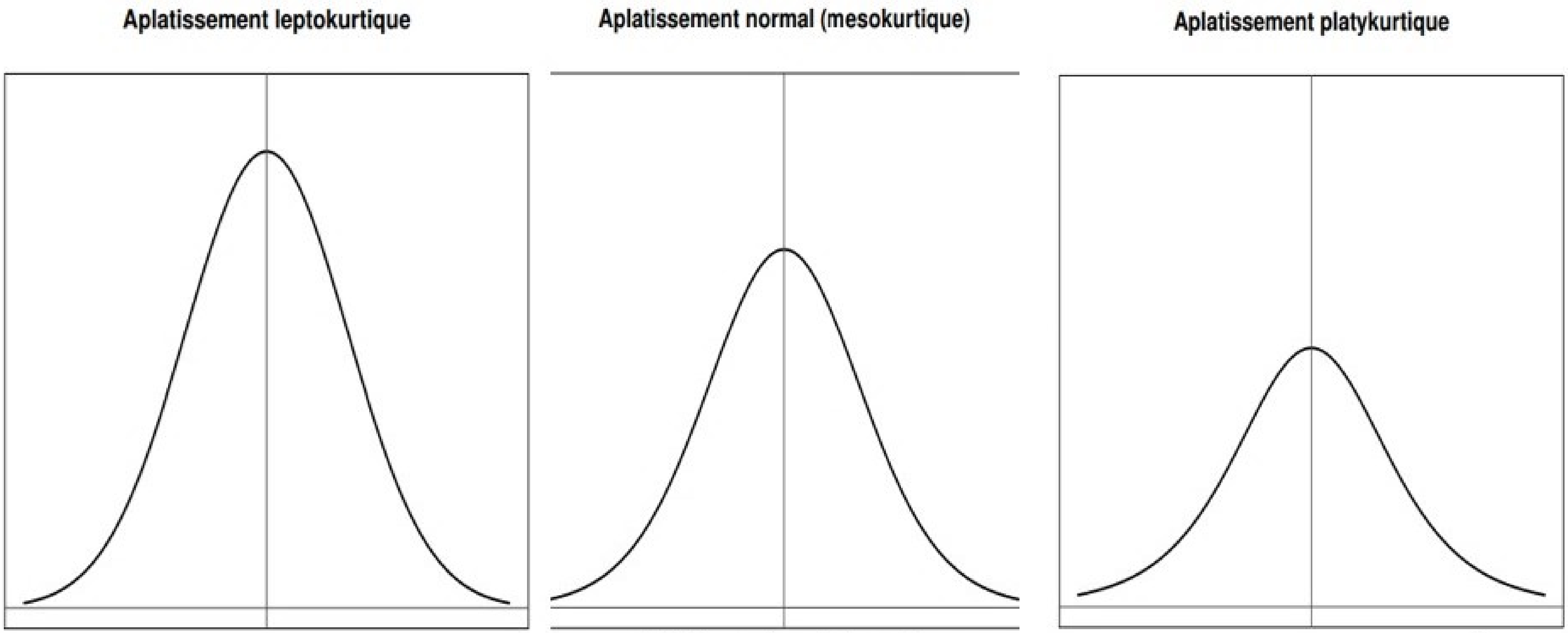
- Il est alors utile de pouvoir mesurer si la forme de la distribution présente une déviation par rapport à l'aplatissement de la distribution normale (***mesokurtique)*** :

  → Une distribution est ***platicurtique*** (ou *hyponormale*) si la courbe est plus aplatie que la courbe normale;

  → Une distribution est ***leptocurtique*** (ou *hypernormale*) si la courbe est plus pointue que la courbe normale.

5. Caractéristiques de forme

L'aplatissement

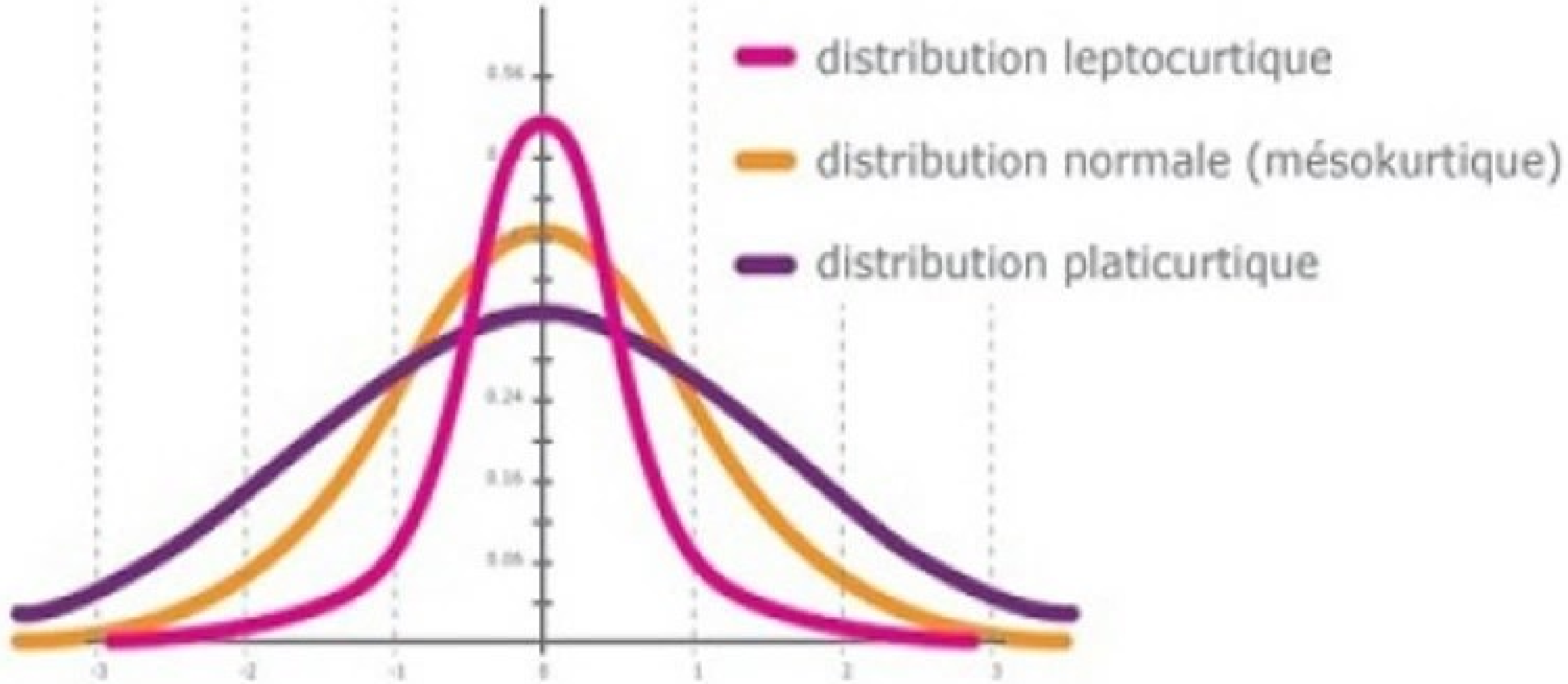


CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

- Comparaison des trois situations :



5. Caractéristiques de forme

L'aplatissement

Coefficient d'aplatissement de Pearson

- Pearson a proposé d'utiliser le coefficient suivant :

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4} \qquad \text{où } \mu_4 = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^4$$

→ $\beta_2 > 3$: Courbe *leptocurtique*, donc plus pointue que la loi normale.

→ $\beta_2 = 3$:Courbe *mesokurtique*, normale .

→ $\beta_2 < 3$: Courbe *platicurtique*, donc plus aplatie que la loi normale.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

Coefficient d'aplatissement de Fisher

- Par analogie avec le coefficient d'asymétrie, Il reste plus naturel de considérer que la valeur de référence est 0 et non pas 3.
- Pour cela, Fisher a proposé d'adopter comme coefficient d'aplatissement la quantité :

$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\sigma^4} - 3$$

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

Coefficient d'aplatissement de Fisher

→$\gamma_2 > 0$: Courbe *leptocurtique,* donc plus pointue que la loi normale.

→$\gamma_2 = 0$:Courbe *mesokurtique*, normale .

→$\gamma_2 < 0$: Courbe *platicurtique*, donc plus aplatie que la loi normale.

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

Exemple

- *On considère la distribution suivante comportant 20 valeurs numériques :*

6.77	7.19	8.40	8.43	9.10	9.21	9.42	9.53	9.75	9.77
9.97	10.43	10.82	11.04	11.13	11.25	11.89	12.03	12.44	13.00

- *Calculer les coefficients d'aplatissement de Pearson et Fisher.*

5. Caractéristiques de forme

L'aplatissement

Exemple: coefficient d'aplatissement de Pearson

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$:

$\bar{\mathbf{x}}$ =*10.07*

$$\mu_4 = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^4$$

$\boldsymbol{\mu_4} \approx \mathbf{17}$

$$\mu_2 = \frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \bar{x})^2$$

$\boldsymbol{\mu_2} \approx \mathbf{2.64}$

$$\boldsymbol{\beta_2} = \frac{17}{2.64^2} = \mathbf{2.44}$$

$\beta_2 < 3$; la distribution est platikurtique (plus aplatie que la loi normale).

CHAPITRE 1 : SÉRIE STATISTIQUE À UNE SEULE VARIABLE

5. Caractéristiques de forme

L'aplatissement

Exemple: coefficient d'aplatissement de Fisher

$$\gamma_2 = \beta_2 - 3 \qquad \boldsymbol{\gamma_2} = 2.44 - 3 \qquad \boldsymbol{\gamma_2 = -0.56}$$

$\gamma_2 < 0$, la distribution est platikurtique (plus aplatie que la loi normale).

RENDEZ VOUS LA SEMAINE PROCHAINE