Pr. Oubrahim
Pr. Madkour

TD de Statistique descriptive

Série 2 : Caractéristiques des distributions à un caractère.

Ex. 1 — **Cours boursiers**

Dans cet exercice, on s'intéresse à la description et à la comparaison des cours boursiers des actions de Maroc Telecom et de France Telecom entre le 17/11/2014 et le 11/12/2014. Les cours sont récapitulés dans les tableaux suivants :

Table 1 – Cours de l'action de Maroc Telecom

Cours en Dh	[115 ; 116 [	[116 ; 116.5 [	[116.5 ; 117 [	[117; 118]
Effectifs	3	8	6	1

Table 2 – Cours de l'action de France Telecom

Cours en €	[13 ; 13.5 [	[13.5 ; 14 [	[14 ; 14.5 [	[14.5; 15]
Effectifs	3	8	5	2

Pour vous guider dans ce travail, on vous demande de calculer quelques indicateurs statistiques :

1. Indicateurs de position :
   a) Calculer le mode de chaque série. Interpréter.
   b) Calculer la moyenne arithmétique de chaque série. Interpréter.
   c) Calculer la médiane de chaque série. Interpréter.
   d) Calculer le premier quartile de chaque série. Interpréter.
   e) Calculer le troisième quartile de chaque série. Interpréter.
2. Indicateurs de dispersion :
   a) Calculer l'étendue de chaque série. Interpréter.
   b) Calculer l'intervalle interquartile de chaque série. Interpréter.
   c) Calculer l'écart-type de chaque série. Interpréter.
   d) Calculer le coefficient de variation de chaque série. Interpréter.
3. Indicateurs de forme :
   a) Calculer le coefficient d'asymétrie de chaque série. Interpréter.
   b) Calculer le coefficient d'aplatissement de chaque série. Interpréter.
4. Quelles sont vos conclusions ?

Answer (Ex. 1) —

1. Indicateurs de position :

a) **Les modes :**

Le mode d'une variable statistique continue se calcule sur la base des effectifs corrigés n_i^c :

Maroc Telecom

Cours en Dh	n_i	a_i	n_i^c
[115 ; 116 [	3	1	3
[116 ; 116.5 [	8	0.5	16
[116.5 ; 117 [	6	0.5	12
[117; 118]	1	1	1

France Telecom

Cours en €	n_i	a_i	n_i^c
[13 ; 13.5 [	3	0.5	6
[13.5 ; 14 [	8	0.5	16
[14; 14.5]	5	0.5	10
[14.5; 15]	2	0.5	4

La classe modale du cours de l'action de Maroc Telecom est la classe ayant le plus grand effectif corrigé, en l'occurrence la classe [116 ; 116.5 [. Le mode est donné par la formule suivante :

$$\begin{aligned} M_o &= 116 + \frac{16-3}{(16-3)+(16-12)} \times (116.5 - 116) \\ &= 116.3824\text{Dh} \end{aligned}$$

Le cours de l'action de Maroc Telecom le plus fréquemment rencontré entre le 17/11/2014 et le 11/12/2014 est de 116.3824 Dh.

De la même manière, la classe modale du cours de l'action de France Telecom est la classe [13.5 ; 14 [. Le cours modal est donné par la formule suivante :

$$\begin{aligned} M_o &= 13.5 + \frac{16-4}{(16-4)+(16-10)} \times (14 - 13.5) \\ &= 13.8333€ \end{aligned}$$

13.8333€ est le cours de l'action de France Telecom le plus fréquemment rencontré entre le 17/11/2014 et le 11/12/2014.

b) **Les moyennes arithmétiques :**

Maroc Telecom

Cours en Dh	n_i	c_i	n_ic_i
[115 ; 116 [	3	115.50	346.50
[116 ; 116.5 [	8	116.25	930.00
[116.5 ; 117 [	6	116.75	700.50
[117; 118]	1	117.50	117.50
Somme	18	–	2094.5

France Telecom

Cours en €	n_i	c_i	n_ic_i
[13 ; 13.5 [	3	13.25	39.75
[13.5 ; 14 [	8	13.75	110
[14; 14.5]	5	14.25	71.25
[14.5; 15]	2	14.75	29.50
Somme	18	–	250.50

Le cours moyen de l'action de Maroc Telecom est :

$$\overline{x} = \frac{1}{18}\sum_{i=1}^{4} n_i c_i = \frac{2094.5}{18} = 116.3611 Dh$$

Les cours de cette action sont répartis autour d'une valeur centrale de 116.3611 Dh.

Le cours moyen de l'action de France Telecom est :

$$\overline{x} = \frac{1}{18}\sum_{i=1}^{4} n_i c_i = \frac{250.50}{18} = 13.9167€$$

Les cours de cette action sont distribués autour d'une valeur centrale de 13.9167 €.

c) **Les médianes :**

La classe médiane correspond à la fréquence cumulée croissante immédiatement supérieure à 0.5 :

Maroc Telecom

Cours en Dh	n_i	f_i	F_i^+
[115 ; 116 [	3	0.17	0.17
[116 ; 116.5 [	8	0.44	0.61
[116.5 ; 117 [	6	0.33	0.94
[117; 118]	1	0.06	1
Somme	18	1	–

France Telecom

Cours en €	n_i	f_i	F_i^+
[13 ; 13.5 [	3	0.17	0.17
[13.5 ; 14 [	8	0.44	0.61
[14; 14.5]	5	0.28	0.89
[14.5; 15]	2	0.11	1
Somme	18	1	–

La classe médiane du cours de l'action de Maroc Telecom est la classe [116 ; 116.5 [. La médiane se calcule par interpolation linéaire comme suit :

$$\begin{aligned} M_e &= 116 + \frac{0.5 - 0.17}{0.61 - 0.17} \times (116.5 - 116) \\ &= 116.3750\text{Dh} \end{aligned}$$

50% environ des cours de l'action de Maroc Telecom sont inférieurs à 116.3750 Dh et 50% environ des cours sont supérieurs à cette valeur.

La classe médiane du cours de l'action de France Telecom est [13.5 ; 14 [. La médiane est donnée par la formule suivante :

$$\begin{aligned} M_e &= 13.5 + \frac{0.5 - 0.17}{0.61 - 0.17} \times (14 - 13.5) \\ &= 13.8750€ \end{aligned}$$

50% environ des cours de l'action de France Telecom sont inférieurs à 13.8750 € et 50% environ des cours sont supérieurs à cette valeur.

d) **Les premiers quartiles :**

La classe contenant le premier quartile correspond à la fréquence cumulée croissante immédiatement supérieure à 0.25. Dans le cas de Maroc Telecom, il s'agit de la classe [116 ; 116.5 [. Le premier quartile se calcule par interpolation linéaire de la façon suivante :

$$\begin{aligned} Q_1 &= 116 + \frac{0.25 - 0.17}{0.61 - 0.17} \times (116.5 - 116) \\ &= 116.0909\text{Dh} \end{aligned}$$

25 % environ des cours de l'action de Maroc Telecom sont inférieurs à 116.0909 Dh et 75% environ des cours sont supérieurs à cette valeur. Dans le cas de France Telecom, la classe qui contient le premier quartile est $[13.5\,;14\,[$. Le premier quartile est donné par :

$$\begin{aligned} Q_1 &= 13.5 + \frac{0.25 - 0.17}{0.61 - 0.17} \times (14 - 13.5) \\ &= 13.5909€ \end{aligned}$$

25 % environ des cours de l'action de France Telecom sont inférieurs à 13.5909€ et 75% environ des cours sont supérieurs à cette valeur.

e) **Les troisièmes quartiles :**
La classe contenant le troisième quartile correspond à la fréquence cumulée croissante immédiatement supérieure à 0.75. Dans le cas de Maroc Telecom, il s'agit de la classe $[116.5\,;117\,[$. Le troisième quartile se calcule par interpolation linéaire de la façon suivante :

$$\begin{aligned} Q_3 &= 116.5 + \frac{0.75 - 0.61}{0.94 - 0.61} \times (117 - 116.5) \\ &= 116.7121\text{Dh} \end{aligned}$$

75 % environ des cours de l'action de Maroc Telecom sont inférieurs à 116.7121 Dh et 25% environ des cours sont supérieurs à cette valeur.
Dans le cas de France Telecom, la classe qui contient le troisième quartile est $[14\,;14.5\,[$. Le troisième quartile est donné par :

$$\begin{aligned} Q_3 &= 14 + \frac{0.75 - 0.61}{0.89 - 0.61} \times (14.5 - 14) \\ &= 14.2500€ \end{aligned}$$

75 % environ des cours de l'action de France Telecom sont inférieurs à 14.2500 € et 25% environ des cours sont supérieurs à cette valeur.

2. Indicateurs de dispersion :

a) **Les étendues :**
Le cours maximal de l'action de Maroc Telecom est de 118 Dh et son cours minimal est de 115 Dh. L'étendue de la série des cours de cette action est :

$$E = 118 - 115 = 3Dh$$

Les cours de l'action de Maroc Telecom fluctuent dans un intervalle de variation de longueur 3 Dh.
Le cours maximal de l'action de France Telecom est de 15 € et son cours minimal est de 12 €.

L'étendue de la série des cours de cette action est :

$$E = 15 - 12 = 3€$$

Les cours de l'action de France Telecom fluctuent dans un intervalle de variation de longueur 3 €.

b) **Les intervalles interquartiles :**
Le premier quartile du cours de l'action de Maroc Telecom est de 116.0909 Dh et son troisième quartile est de 116.7121 Dh. L'intervalle interquartile correspondant est :

$$IQ_{0.25} = 116.7121 - 116.0909 = 0.6212Dh$$

50% des cours de cette action se trouvent entre 116.0909 Dh et 116.7121 Dh.
Le premier quartile du cours de l'action de France Telecom est de 13.5909 € et son troisième quartile est de 14.2500 €. L'intervalle interquartile correspondant est :

$$IQ_{0.25} = 14.2500 - 13.5909 = 0.6591€$$

50% des cours de cette action sont compris entre 13.5909 € et 14.2500 €.

c) **Les écarts-types :**
La variance se calcule à l'aide de deux formules. On va utiliser une formule pour les cours de l'action de Maroc Telecom et l'autre pour les cours de l'action de France Telecom :

Table 3 – Maroc Telecom

Cours en Dh	n_i	c_i	c_i^2	$n_i c_i^2$
[115 ; 116 [	3	115.50	13340.25	40020.75
[116 ; 116.5 [	8	116.25	13514.06	108112.5
[116.5 ; 117 [	6	116.75	13630.56	81783.36
[117; 118]	1	117.50	13806.25	13806.25
Somme	18	–	–	243722.9

La variance des cours de Maroc Telecom est donnée par :

$$\mathbb{V}(X) = \overline{c^2} - \overline{x}^2 = \frac{243722.9}{18} - (116.36)^2 = 0.2555Dh^2$$

L'écart-type correspondant est :

$$\sigma = \sqrt{0.2555} = 0.5055Dh$$

Les cours de cette action s'éloignent de 0.5055 Dh du cours moyen $\overline{x} = 116.3611$ Dh.

TABLE 4 – France Telecom

Cours en €	n_i	c_i	$c_i - \overline{x}$	$(c_i - \overline{x})^2$	$n_i\,(c_i - \overline{x})^2$
[13 ; 13.5 [	3	13.25	-0.67	0.4489	1.3467
[13.5 ; 14 [	8	13.75	-0.17	0.0289	0.2312
[14; 14.5]	5	14.25	0.33	0.1089	0.5445
[14.5; 15]	2	14.75	0.83	0.6889	1.3778
Somme	18	–	–	–	3.5002

La variance des cours de France Telecom est donnée par :

$$\mathbb{V}\,(X) = \frac{1}{18}\sum_{i=1}^{4} n_i\,(c_i - \overline{x})^2 = \frac{3.5002}{18} = 0.1945€^2$$

L'écart-type correspondant est :

$$\sigma = \sqrt{0.1945} = 0.4410€$$

Les cours de cette action s'éloignent de 0.4410 € du cours moyen $\overline{x} = 13.9167$ €.

d) **Les coefficients de variation :**

Le coefficient de variation des cours de l'action de Maroc Telecom est :

$$CV = \frac{\sigma}{\overline{x}} = \frac{0.5055}{116.3611} = 0.0043$$

La dispersion relative des cours de cette action est de 0.0043.

Le coefficient de variation des cours de l'action de France Telecom est :

$$CV = \frac{\sigma}{\overline{x}} = \frac{0.4410}{13.9167} = 0.0317$$

La dispersion relative des cours de cette action est de 0.0317.

3. INDICATEURS DE FORME :

a) **Les coefficients d'asymétrie :**

TABLE 5 – Maroc Telecom

Cours en Dh	n_i	c_i	$\frac{c_i-\overline{x}}{\sigma}$	$\left(\frac{c_i-\overline{x}}{\sigma}\right)^3$	$n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^3$
[115 ; 116 [	3	115.50	-1.2113	-1.7773	-5.3319
[116 ; 116.5 [	8	116.25	-0.1549	-0.0037	-0.0296
[116.5; 117]	6	116.75	0.5493	0.1657	0.9942
[117; 118]	1	117.50	1.6056	4.1392	4.1392
Somme	18	–	–	–	-0.2281

Le coefficient d'asymétrie de Fisher de la distribution des cours de l'action de Maroc Telecom

Table 6 – France Telecom

Cours en €	n_i	c_i	$\frac{c_i-\overline{x}}{\sigma}$	$\left(\frac{c_i-\overline{x}}{\sigma}\right)^3$	$n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^3$
[13 ; 13.5 [	3	13.25	-1.5227	-3.5306	-10.5918
[13.5 ; 14 [	8	13.75	-0.3864	-0.0577	-0.4616
[14; 14.5]	5	14.25	0.7500	0.4219	2.1095
[14.5; 15]	2	14.75	1.8864	6.7128	13.4256
Somme	18	–	–	–	4.4817

est :

$$\gamma_1 = \frac{1}{18}\sum_{i=1}^{4} n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^3 = \frac{-0.2281}{18} = -0.0127$$

Le coefficient d'asymétrie étant légèrement inférieur à 0, la distribution des cours de cette action est légèrement asymétrique à gauche. Les petits cours se produisent légèrement plus que les cours élevés.

Le coefficient d'asymétrie de Fisher de la distribution des cours de l'action de France Telecom est :

$$\gamma_1 = \frac{1}{18}\sum_{i=1}^{4} n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^3 = \frac{4.4817}{18} = 0.2490$$

Le coefficient d'asymétrie étant supérieur à 0, la distribution des cours de cette action est asymétrique à droite. Les cours élevés se produisent plus que les petits cours.

b) **Les coefficients d'aplatissement :**

Table 7 – Maroc Telecom

Cours en Dh	n_i	c_i	$\frac{c_i-\overline{x}}{\sigma}$	$\left(\frac{c_i-\overline{x}}{\sigma}\right)^4$	$n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^4$
[115 ; 116 [	3	115.50	-1.2113	2.1528	6.4584
[116 ; 116.5 [	8	116.25	-0.1549	0.0006	0.0048
[116.5; 117]	6	116.75	0.5493	0.0910	0.5460
[117; 118]	1	117.50	1.6056	6.6458	6.6458
Somme	18	–	–	–	13.6550

Table 8 – France Telecom

Cours en €	n_i	c_i	$\frac{c_i-\overline{x}}{\sigma}$	$\left(\frac{c_i-\overline{x}}{\sigma}\right)^4$	$n_i\left(\frac{c_i-\overline{x}}{\sigma}\right)^4$
[13 ; 13.5 [	3	13.25	-1.5227	5.3760	16.1280
[13.5 ; 14 [	8	13.75	-0.3864	0.0223	0.1784
[14; 14.5]	5	14.25	0.7500	0.3164	1.5820
[14.5; 15]	2	14.75	1.8864	12.6630	25.3260
Somme	18	–	–	–	43.2144

Le coefficient d'aplatissement de Pearson de la distribution des cours de l'action de Maroc Telecom est :

$$\beta_2 = \frac{1}{18}\sum_{i=1}^{4} n_i \left(\frac{c_i - \overline{x}}{\sigma}\right)^4 = \frac{13.6550}{18} = 0.7586$$

Le coefficient d'aplatissement étant inférieur à 3, la distribution des cours de l'action de Maroc Telecom est plus aplatie que la normale. Les cours extrêmes sont très fréquents.

Le coefficient d'aplatissement de Pearson de la distribution des cours de l'action de France Telecom est :

$$\beta_2 = \frac{1}{18}\sum_{i=1}^{4} n_i \left(\frac{c_i - \overline{x}}{\sigma}\right)^4 = \frac{43.2144}{18} = 2.4008$$

Le coefficient d'aplatissement étant inférieur à 3, la distribution des cours de l'action de France Telecom est plus aplatie que la normale. Les cours extrêmes sont très fréquents.

4. CONCLUSIONS

L'ensemble des indicateurs calculés dans les questions précédentes pour les cours des actions de Maroc Telecom et de France Telecom sont regroupés dans le tableau suivant :

Indicateurs		Maroc Telecom	France Telecom
POSITION	M_o	116.3824 Dh	13.8333 €
	$\overline{x}$	116.3611 Dh	13.9167 €
	M_e	116.3750 Dh	13.8750 €
	Q_1	116.0909 Dh	13.5909 €
	Q_3	116.7121 Dh	14.2500 €
DISPERSION	E	3 Dh	3 €
	$IQ_{0.25}$	0.6212 Dh	0.6591 €
	σ	0.5055 Dh	0.4410 €
	CV	0.0043	0.0317
FORME	γ_1	-0.0127	0.2490
	β_2	0.7586	2.4008

a) **Description des cours de l'action de Maroc Telecom**

Les cours de l'action de Maroc Telecom varient entre 115 Dh et 118 Dh, il fluctent donc dans un intervalle de longueur $E = 3$ Dh et se répartissent autour d'un centre, mesuré par le mode, la médiane et la moyenne arithmétique, qui si situe au voisinage de 116.375 Dh. Les cours s'éloignent en moyenne de $\sigma = 0.5055$ Dh du cours moyen $\overline{x} = 116.3611$ Dh. 50% des cours sont inférieurs au cours médian ($M_e = 116.3750$ Dh) et 50% des cours sont supérieurs à ce cours médian. 50% des cours sont compris entre le premier quartile ($Q_1 = 116.0909$ Dh) et le troisième quartile ($Q_3 = 116.7121$ Dh). Les trois mesures de centre sont ordonnées dans le sens $\overline{x} < M_e < M_o$, ceci est le signe d'une asymétrie à gauche qui est confirmée par la valeur négative du coefficient d'asymétrie de Fisher ($\gamma_1 = -0.0127$). Par conséquent, les cours qui sont inférieurs au centre sont plus nombreux que les cours qui lui sont supérieurs. Le coefficient d'aplatissement de Pearson étant inférieur à 3 ($\beta_2 = 0.7586$), il indique que la distribution des cours de l'action de Maroc Telecom est plus aplatie que la normale, cela veut dire que les cours extrêmes (les plus faibles et les plus élevés) apparaissent de façon anormalement élevée.

b) **Description des cours de l'action de France Telecom**
Les cours de l'action de France Telecom varient entre 13 € et 15 €, il fluctent donc dans un intervalle de longueur $E = 3$ € et se répartissent autour d'un centre, mesuré par le mode, la médiane et la moyenne arithmétique, qui si situe au voisinage de 13.875 €. Les cours s'éloignent en moyenne de $\sigma = 0.4410$ € du cours moyen $\overline{x} = 13.9167$ €. 50% des cours sont inférieurs au cours médian ($M_e = 13.8750$ €) et 50% des cours sont supérieurs à ce cours médian. 50% des cours sont compris entre le premier quartile ($Q_1 = 13.5909$ €) et le troisième quartile ($Q_3 = 14.2500$ €). Les trois mesures de centre sont ordonnées dans le sens $M_o < M_e < \overline{x}$, ceci est le signe d'une asymétrie à droite qui est confirmée par la valeur positive du coefficient d'asymétrie de Fisher ($\gamma_1 = 0.2490$). Par conséquent, les cours qui sont supérieurs au centre sont plus nombreux que les cours qui lui sont inférieurs. Le coefficient d'aplatissement de Pearson étant inférieur à 3 ($\beta_2 = 2.4008$), il indique que la distribution des cours de l'action de France Telecom est plus aplatie que la normale, cela veut dire que les cours extrêmes (les plus faibles et les plus élevés) apparaissent de façon anormalement élevée.

c) **Comparaison des cours des actions de Maroc Telecom et de France Telecom**
Les cours des actions de Maroc Telecom et de France Telecom n'étant pas exprimés dans la même monnaie, on ne peut donc comparer leur distributions respectives que sur la base d'indicateurs sans dimension tels que le coefficient de variation, le coefficient d'asymétrie et le coefficient d'aplatissement :

- Dispersion relative : le coefficient de variation des cours de l'action de France Telecom ($CV = 0.0317$) est supérieur à celui des cours de l'action de Maroc Telecom ($CV = 0.0043$). Ceci signifie que les cours de l'action France Telecom sont plus volatiles que dans le cas de Maroc Telecom.
- Asymétrie : les distributions des cours des action étudiées sont asymétriques. Néanmoins, l'asymétrie à droite de la distributions des cours de l'action de France Telecom est plus prononcée que l'asymétrie à gauche de la distribution des cours de l'action de Maroc Telecom qui peut être négligée [1].
- Aplatissement : les deux distributions des cours étudiés sont plus aplaties que la normale. Cependant, la distribution des cours de l'action de Maroc Telecom ($\beta_2 = 0.7586$) est plus aplatie que dans le cas de France Telecom ($\beta_2 = 2.4008$). Concrètement, cela veut dire que les cours extrêmes sont plus fréquents dans le cas de Maroc Telecom que dans le cas de France Telecom.

Ex. 2 — Emploi par branche d'activité
Le tableau suivant présente le nombre annuel (en milliers) d'emplois créés par deux branches d'activité au niveau national entre 2004 et 2013. Les données sont issues du site du Haut Commissariat au Plan.

1. Un test statistique de l'asymétrie d'une distribution permet de savoir si cette asymétrie est néligeable ou si, au contraire, elle est significative.

TABLE 9 – Emploi annuel par branche d'activité au niveau national

Années	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Services	34.7	35.1	36.3	36.7	37.2	37.6	37.5	38.5	39.3	39.7
Industrie	19.5	19.6	20.3	21.1	21.7	21.7	22.1	21.8	21.4	20.8

On vous demande de décrire et de comparer les deux séries de données, et ce, en calculant certains indicateurs statistiques.

1. Indicateurs de position :
   a) Calculer la moyenne arithmétique de chaque série. Interpréter.
   b) Calculer la médiane de chaque série. Interpréter.
   c) Calculer le premier quartile de chaque série. Interpréter.
   d) Calculer le troisième quartile de chaque série. Interpréter.
2. Indicateurs de dispersion :
   a) Calculer l'étendue de chaque série. Interpréter.
   b) Calculer l'intervalle interquartile de chaque série. Interpréter.
   c) Calculer l'écart-type de chaque série. Interpréter.
   d) Calculer le coefficient de variation de chaque série. Interpréter.
3. Indicateurs de forme :
   a) Calculer le coefficient d'asymétrie de chaque série. Interpréter.
   b) Calculer le coefficient d'aplatissement de chaque série. Interpréter.
4. Quelles sont vos conclusions ?

Answer (Ex. 2) —

1. INDICATEURS DE POSITION :
   a) **Les moyennes arithmétiques :**

Services

x_i (en milliers)	n_i	$n_i x_i$
34.7	1	34.7
35.1	1	35.1
36.3	1	36.3
36.7	1	36.7
37.2	1	37.2
37.5	1	37.5
37.6	1	37.6
38.5	1	38.5
39.3	1	39.3
39.7	1	39.7
Somme	10	372.6

Industrie

x_i (en milliers)	n_i	$n_i x_i$
19.5	1	19.5
19.6	1	19.6
20.3	1	20.3
20.8	1	20.8
21.1	1	21.1
21.4	1	21.4
21.7	2	43.4
21.8	1	21.8
22.1	1	22.1
Somme	10	210

Le nombre moyen d'emplois créés annuellement par les services au niveau national est :

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{10} n_i x_i = \frac{372.6}{10} = 37.26 \textit{ mille emplois}$$

Les nombres d'emplois créés annuellement par les services au niveau national sont répartis autour d'une valeur centrale de 37260 emplois.
Le nombre moyen d'emplois créés annuellement par l'industrie au niveau national est :

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{9} n_i c_i = \frac{210}{10} = 21 \textit{ mille emplois}$$

Les nombres d'emplois créés annuellement par l'industrie au niveau national sont répartis autour d'une valeur centrale de 21000 emplois.

b) **Les médianes :**

Services

x_i (en milliers)	n_i	f_i	F_i^+
34.7	1	0.1	0.1
35.1	1	0.1	0.2
36.3	1	0.1	0.3
36.7	1	0.1	0.4
37.2	1	0.1	0.5
37.5	1	0.1	0.6
37.6	1	0.1	0.7
38.5	1	0.1	0.8
39.3	1	0.1	0.9
39.7	1	0.1	1
Somme	10	1	–

Industrie

x_i (en milliers)	n_i	f_i	F_i^+
19.5	1	0.1	0.1
19.6	1	0.1	0.2
20.3	1	0.1	0.3
20.8	1	0.1	0.4
21.1	1	0.1	0.5
21.4	1	0.1	0.6
21.7	2	0.2	0.8
21.8	1	0.1	0.9
22.1	1	0.1	1
Somme	10	1	–

Le nombre médian d'emplois créés annuellement par les services au niveau national appartient à l'intervalle médian $[37.2; 37.5]$. Il est égal à :

$$\begin{aligned} M_e &= \frac{37.2 + 37.5}{2} \\ &= 37.35 \textit{ mille emplois} \end{aligned}$$

50% environ des nombres d'emplois créés par les services sont inférieurs à 37350 emplois et 50% environ des nombres d'emplois créés sont supérieurs à cette valeur.

Le nombre médian d'emplois créés annuellement par l'industrie au niveau national appartient à l'intervalle médian $[21.1; 21.4]$. Il est égal à :

$$\begin{aligned} M_e &= \frac{21.1 + 21.4}{2} \\ &= 21.25 \textit{ mille emplois} \end{aligned}$$

50% environ des nombres d'emplois créés par l'industrie sont inférieurs à 21250 emplois et 50% environ des nombres d'emplois créés sont supérieurs à cette valeur.

c) **Les premiers quartiles :**

Le premier quartile correspond à la fréquence cumulée croissante immédiatement supérieure à 0.25. Dans le cas des services, on a :

$$Q_1 = 36.3 \textit{ mille emplois}$$

25% environ des nombres d'emplois créés par les services sont inférieurs à 36300 emplois et 75% environ des nombres d'emplois créés sont supérieurs à cette valeur.

Dans le cas de l'industrie, le premier quartile est égal à :

$$Q_1 = 20.3 \ mille \ emplois$$

25% environ des nombres d'emplois créés par l'industrie sont inférieurs à 20300 emplois et 75% environ des nombres d'emplois créés sont supérieurs à cette valeur.

d) **Les troisièmes quartiles :**
Le troisième quartile correspond à la fréquence cumulée croissante immédiatement supérieure à 0.75. Dans le cas des services, on a :

$$Q_3 = 38.5 \ mille \ emplois$$

75% environ des nombres d'emplois créés par les services sont inférieurs à 38500 emplois et 25% environ des nombres d'emplois créés sont supérieurs à cette valeur.
Dans le cas de l'industrie, le troisième quartile est égal à :

$$Q_3 = 21.7 \ mille \ emplois$$

75% environ des nombres d'emplois créés par l'industrie sont inférieurs à 21700 emplois et 25% environ des nombres d'emplois créés sont supérieurs à cette valeur.

2. Indicateurs de dispersion :

a) **Les étendues :**
Le nombre maximal d'emplois créés par les services est de 39700 et le nombre minimal d'emplois créés par les services est de 34700. L'étendue de la série des nombres d'emplois créés par cette branche est :

$$E = 39700 - 34700 = 5000 \ emplois$$

Les nombres d'emplois créés annuellement par les services fluctuent dans un intervalle de variation de longueur 5000 emplois.
Le nombre maximal d'emplois créés par l'industrie est de 22100 et le nombre minimal d'emplois créés par l'industrie est de 19500. L'étendue de la série des nombres d'emplois créés par cette branche est :

$$E = 22100 - 19500 = 2600 \ emplois$$

Les nombres d'emplois créés annuellement par l'industrie fluctuent dans un intervalle de variation de longueur 2600 emplois.

b) **Les intervalles interquartiles :**
Le premier quartile des nombres d'emplois créés annuellement par les services est égal à 36300 emplois et le troisième quartile est égal à 38500. L'intervalle interquartile correspondant est :

$$IQ_{0.25} = 38500 - 36300 = 2200 \ emplois$$

50% des nombres d'emplois créés annuellement pas les services sont compris entre 36300 et 38500.
Le premier quartile des nombres d'emplois créés annuellement par l'industrie est égal à 20300

emplois et le troisième quartile est égal à 21700. L'intervalle interquartile correspondant est :

$$IQ_{0.25} = 21700 - 20300 = 1400 \; emplois$$

50% des nombres d'emplois créés annuellement pas l'industrie sont compris entre 20300 et 21700.

c) **Les écarts-types :**
La variance se calcule à l'aide de deux formules. On va utiliser une formule pour les nombres d'emplois créés par les services et l'autre pour les nombres d'emplois créés par l'industrie :

TABLE 10 – Services

x_i (en millers)	n_i	x_i^2	$n_i x_i^2$
34.7	1	1204.09	1204.09
35.1	1	1232.01	1232.01
36.3	1	1317.69	1317.69
36.7	1	1346.89	1346.89
37.2	1	1383.84	1383.84
37.5	1	1406.25	1406.25
37.6	1	1413.76	1413.76
38.5	1	1482.25	1482.25
39.3	1	1544.49	1544.49
39.7	1	1576.09	1576.09
Somme	10	–	13907.36

La variance des nombres d'emplois créés annuellement par les services est donnée par :

$$\mathbb{V}(X) = \overline{x^2} - \overline{x}^2 = \frac{13907.36}{10} - (37.26)^2 = 2.4284 \, (mille \; emplois)^2$$

L'écart-type correspondant est :

$$\sigma = \sqrt{2.4284} = 1.558332 \; mille \; emplois$$

Les nombres d'emplois créés annuellement par les services s'éloignent de 1558 emplois du nombre d'emplois moyen 37260.
La variance des nombres d'emplois créés annuellement par l'industrie est donnée par :

$$\mathbb{V}(X) = \frac{1}{10} \sum_{i=1}^{9} n_i \, (x_i - \overline{x})^2 = \frac{7.74}{10} = 0.774 \, (mille \; emplois)^2$$

L'écart-type correspondant est :

$$\sigma = \sqrt{0.774} = 0.8797727 \; mille \; emplois$$

TABLE 11 – Industrie

x_i (en millers)	n_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$n_i\,(x_i - \overline{x})^2$
19.5	1	-1.5	2.25	2.25
19.6	1	-1.4	1.96	1.96
20.3	1	-0.7	0.49	0.49
20.8	1	-0.2	0.04	0.04
21.1	1	0.1	0.01	0.01
21.4	1	0.4	0.16	0.16
21.7	2	0.7	0.49	0.98
21.8	1	0.8	0.64	0.64
22.1	1	1.1	1.21	1.21
Somme	10	–	–	7.74

Les nombres d'emplois créés annuellement par l'industrie s'éloignent de 880 emplois du nombre d'emplois moyen 21000.

d) **Les coefficients de variation :**

Le coefficient de variation des nombres d'emplois créés annuellement par les services est :

$$CV = \frac{\sigma}{\overline{x}} = \frac{1.558332}{37.26} = 0.0418$$

La dispersion relative des nombres d'emplois créés annuellement par les services est de 0.0418.
Le coefficient de variation des nombres d'emplois créés annuellement par l'industrie est :

$$CV = \frac{\sigma}{\overline{x}} = \frac{0.8797727}{21} = 0.0419$$

La dispersion relative des nombres d'emplois créés annuellement par l'industrie est de 0.0419.

3. INDICATEURS DE FORME :

a) **Les coefficients d'asymétrie :**

Le coefficient d'asymétrie de Fisher de la distribution des nombres d'emplois créés annuellement par les services est :

$$\gamma_1 = \frac{1}{10}\sum_{i=1}^{10} n_i \left(\frac{x_i - \overline{x}}{\sigma}\right)^3 = \frac{-0.7767}{10} = -0.07767$$

Le coefficient d'asymétrie étant légèrement inférieur à 0, la distribution des nombres d'emplois créés annuellement par les services est légèrement asymétrique à gauche. Les petits nombres d'emplois créés se produisent légèrement plus que les nombres d'emplois élevés.
Le coefficient d'asymétrie de Fisher de la distribution des nombres d'emplois créés annuellement

TABLE 12 – Services

x_i (en millers)	n_i	$\frac{x_i-\overline{x}}{\sigma}$	$\left(\frac{x_i-\overline{x}}{\sigma}\right)^3$	$n_i\left(\frac{x_i-\overline{x}}{\sigma}\right)^3$
34.7	1	-1.6428	-4.4334	-4.4334
35.1	1	-1.3861	-2.6631	-2.6631
36.3	1	-0.6160	-0.2338	-0.2338
36.7	1	-0.3594	-0.0464	-0.0464
37.2	1	-0.0385	-0.0001	-0.0001
37.5	1	0.1540	0.0037	0.0037
37.6	1	0.2182	0.0104	0.0104
38.5	1	0.7957	0.5038	0.5038
39.3	1	1.3091	2.2434	2.2434
39.7	1	1.5658	3.8387	3.8387
Somme	10	–	–	-0.7767

TABLE 13 – Industrie

x_i (en millers)	n_i	$\frac{x_i-\overline{x}}{\sigma}$	$\left(\frac{x_i-\overline{x}}{\sigma}\right)^3$	$n_i\left(\frac{x_i-\overline{x}}{\sigma}\right)^3$
19.5	1	-1.7050	-4.9564	-4.9564
19.6	1	-1.5913	-4.0297	-4.0297
20.3	1	-0.7957	-0.5037	-0.5037
20.8	1	-0.2273	-0.0117	-0.0117
21.1	1	0.1137	0.0015	0.0015
21.4	1	0.4547	0.0940	0.0940
21.7	2	0.7957	0.5037	1.0074
21.8	1	0.9093	0.7519	0.7519
22.1	1	1.2503	1.9546	1.9546
Somme	10	–	–	-5.6921

par l'industrie est :

$$\gamma_1 = \frac{1}{10}\sum_{i=1}^{9} n_i \left(\frac{x_i-\overline{x}}{\sigma}\right)^3 = \frac{-5.6921}{10} = -0.56921$$

Le coefficient d'asymétrie étant inférieur à 0, la distribution des nombres d'emplois créés annuellement par l'industrie est asymétrique à gauche. Les petits nombres d'emplois créés se produisent plus que les nombres d'emplois élevés.

b) **Les coefficients d'aplatissement :**

Le coefficient d'aplatissement de Pearson de la distribution des nombres d'emplois créés annuel-

Table 14 – Services

x_i (en millers)	n_i	$\frac{x_i-\overline{x}}{\sigma}$	$\left(\frac{x_i-\overline{x}}{\sigma}\right)^4$	$n_i\left(\frac{x_i-\overline{x}}{\sigma}\right)^4$
34.7	1	-1,6428	7,2832	7,2832
35.1	1	-1,3861	3,6913	3,6913
36.3	1	-0,6160	0,1440	0,1440
36.7	1	-0,3594	0,0167	0,0167
37.2	1	-0,0385	0,0000	0,0000
37.5	1	0,1540	0,0006	0,0006
37.6	1	0,2182	0,0023	0,0023
38.5	1	0,7957	0,4009	0,4009
39.3	1	1,3091	2,9368	2,9368
39.7	1	1,5658	6,0106	6,0106
Somme	10	–	–	20,4863

Table 15 – Industrie

x_i (en millers)	n_i	$\frac{x_i-\overline{x}}{\sigma}$	$\left(\frac{x_i-\overline{x}}{\sigma}\right)^4$	$n_i\left(\frac{x_i-\overline{x}}{\sigma}\right)^4$
19.5	1	-1.7050	8.4505	8.4505
19.6	1	-1.5913	6.4125	6.4125
20.3	1	-0.7957	0.4008	0.4008
20.8	1	-0.2273	0.0027	0.0027
21.1	1	0.1137	0.0002	0.0002
21.4	1	0.4547	0.0427	0.0427
21.7	2	0.7957	0.4008	0.8016
21.8	1	0.9093	0.6837	0.6837
22.1	1	1.2503	2.4439	2.4439
Somme	10	–	–	19.2386

lement par les services est :

$$\beta_2 = \frac{1}{10}\sum_{i=1}^{10} n_i \left(\frac{x_i-\overline{x}}{\sigma}\right)^4 = \frac{20.4863}{10} = 2.04863$$

Le coefficient d'aplatissement étant inférieur à 3, la distribution des nombres d'emplois créés annuellement par les services est plus aplatie que la normale. Les nombres d'emplois extrêmes sont très fréquents.

Le coefficient d'aplatissement de Pearson de la distribution des nombres d'emplois créés annuellement par l'industrie est :

$$\beta_2 = \frac{1}{10}\sum_{i=1}^{9} n_i \left(\frac{x_i-\overline{x}}{\sigma}\right)^4 = \frac{19.2386}{10} = 1.92386$$

Le coefficient d'aplatissement étant inférieur à 3, la distribution des nombres d'emplois créés annuellement par l'industrie est plus aplatie que la normale. Les nombres d'emplois extrêmes sont très fréquents.

4. CONCLUSIONS

L'ensemble des indicateurs calculés dans les questions précédentes pour les nombres d'emplois créés annuellement par les services et par l'industrie sont regroupés dans le tableau suivant :

Indicateurs		Services	Industrie
POSITION	$\overline{x}$	37260 emplois	21000 emplois
	M_e	37350 emplois	21250 emplois
	Q_1	36300 emplois	20300 emplois
	Q_3	38500 emplois	21700 emplois
DISPERSION	E	5000 emplois	2600 emplois
	$IQ_{0.25}$	2200 emplois	1400 emplois
	σ	1558 emplois	880 emplois
	CV	0.0418	0.0419
FORME	γ_1	-0.07767	-0.56921
	β_2	2.04863	1.92386

a) **Description des nombres d'emplois créés par le secteur tertiaire**

Les nombres d'emplois créés annuellement par les services au niveau national entre 2004 et 2013 varient entre 34700 et 39700 emplois, ils fluctuent donc dans un intervalle d'amplitude $E = 5000$ emplois et se répartissent autour d'un centre, mesuré par la médiane et par la moyenne arithmétique, qui se situe au voisinage de 37350 emplois. Les nombres d'emplois s'éloignent en moyenne de $\sigma = 1558$ emplois du nombre moyen d'emplois créés annuellement, soit $\overline{x} = 37260$ emplois. 50% des nombres d'emplois sont inférieurs au nombre médian d'emplois créés ($M_e = 37350$ emplois) et 50% des nombres d'emplois créés sont supérieurs à ce nombre médian. 50% des nombres d'emplois sont compris entre le premier quartile ($Q_1 = 36300$ emplois) et le troisième quartile ($Q_3 = 38500$ emplois). Les deux mesures de centre sont ordonnées dans le sens $\overline{x} < M_e$, ceci est le signe d'une asymétrie à gauche qui est confirmée par la valeur négative du coefficient d'asymétrie de Fisher ($\gamma_1 = -0.07767$). Par conséquent, les nombres d'emplois créés qui sont inférieurs au centre sont plus nombreux que les nombres d'emplois créés qui lui sont supérieurs. Le coefficient d'aplatissement de Pearson est inférieur à 3 ($\beta_2 = 2.04863$), il indique que la distribution des nombres d'emplois créés est plus aplatie que la normale, cela veut dire que les nombres extrêmes d'emplois créés (les plus faibles et les plus élevés) apparaissent de façon anormalement élevée.

b) **Description des nombres d'emplois créés par le secteur industriel**

Les nombres d'emplois créés annuellement par l'industrie au niveau national entre 2004 et 2013 varient entre 19500 et 22100 emplois, ils fluctuent donc dans un intervalle d'amplitude $E = 2600$ emplois et se répartissent autour d'un centre, mesuré par la médiane et par la moyenne arithmétique, qui se situe au voisinage de 21250 emplois. Les nombres d'emplois s'éloignent en moyenne de $\sigma = 880$ emplois du nombre moyen d'emplois créés annuellement $\overline{x} = 21000$ emplois. 50% des nombres d'emplois sont inférieurs au nombre médian d'emplois créés ($M_e = 21250$ emplois) et 50% des nombres d'emplois créés sont supérieurs à ce nombre médian. 50% des nombres d'emplois sont compris entre le premier quartile ($Q_1 = 20300$ emplois) et le troisième quartile ($Q_3 = 21700$ emplois). Les deux mesures de centre sont ordonnées dans le sens $\overline{x} < M_e$, ceci est le signe d'une asymétrie à gauche qui est confirmée par la valeur négative du coefficient d'asymétrie de Fisher ($\gamma_1 = -0.56921$). Par conséquent, les nombres d'emplois créés qui sont inférieurs au centre sont plus nombreux que les nombres d'emplois créés qui lui sont supérieurs. Le coefficient d'aplatissement de Pearson est inférieur à 3 ($\beta_2 = 1.92386$), il indique que la distribution des nombres d'emplois créés est plus aplatie que la normale, cela veut dire que les nombres extrêmes d'emplois créés (les plus faibles et les plus élevés) apparaissent de façon anormalement élevée.

c) **Comparaison des nombres d'emplois créés par le secteur tertiaire et par le secteur industriel**

Les services ont créés annuellement des nombres d'emplois supérieurs, en moyenne, à ceux créés par l'industrie au niveau national entre 2004 et 2013. Les deux séries des données ont pratiquement le même coefficient de variation (0.0418 et 0.0419 respectivement), elles ont donc le même niveau de dispersion. Le coefficient d'asymétrie de Fisher relatif à l'industrie est plus petit que celui des services, ce qui signifie que la distribution des nombres d'emplois créés par les services est plus asymétrique à gauche que celle des nombres d'emplois créés par l'industrie, c'est à dire que les nombres d'emplois inférieurs au nombre moyen sont plus nombreux dans le cas de l'industrie que dans le cas des services. Enfin, on note que les coefficients d'aplatissement de Pearson sont inférieurs à 3, on conclut que les distributions des données sont plus aplaties que la normale. Ceci signifie que les nombres d'emplois extrêmes sont anormalement élevés et ils sont encore plus élevés dans le cas de l'industrie que dans le cas des services.