

Module 3

Coefficient de variation

Cote z

Quantiles

Élaboré par Afef Ben Zine El Abidine

Formule du coefficient de variation

Le coefficient de variation est une mesure de dispersion relative des données; on le calcule ainsi :

Coefficient de variation
$CV = \frac{\text{écart type}}{\text{moyenne}} \times 100 \% = \frac{\sigma}{\mu} \times 100 \%$

Plus le coefficient de variation est faible, plus les données sont homogènes.

Si $CV < 15\%$, on dit qu'il y a une bonne homogénéité des données.

NOTE

Si l'on travaille avec un échantillon, on remplace σ par s et μ par \bar{x} dans la formule.

Exemple pour illustrer le coefficient de variation (CV):

Salaires, 1930: 30 \$ 37 \$ 44 \$ 50 \$ 55 \$ 60 \$

Moyenne : 46 \$

Écart type : 10,30 \$

Supposons qu'en 2012, le salaire moyen de six couturières (bien entendu, ce ne sont plus celles de 1930 !) était de 356 \$ par semaine, avec un écart type de 10,30 \$. Voici la distribution des salaires des couturières :

Salaires, 2012: 340 \$ 347 \$ 354 \$ 360 \$ 365 \$ 370 \$

Moyenne : 356 \$

Écart type : 10,30 \$

Comment peut-on interpréter ces données? Est-ce que les salaires en 2012 sont plus homogènes qu'en 1930?

Le coefficient de variation (= écart type/moyenne)

Voici le coefficient de variation des salaires des couturières :

$$\text{En 1930: CV} = \frac{10,30 \$}{46 \$} \times 100 \% = 22,4 \%$$

$$\text{En 2012: CV} = \frac{10,30 \$}{356 \$} \times 100 \% = 2,9 \%$$

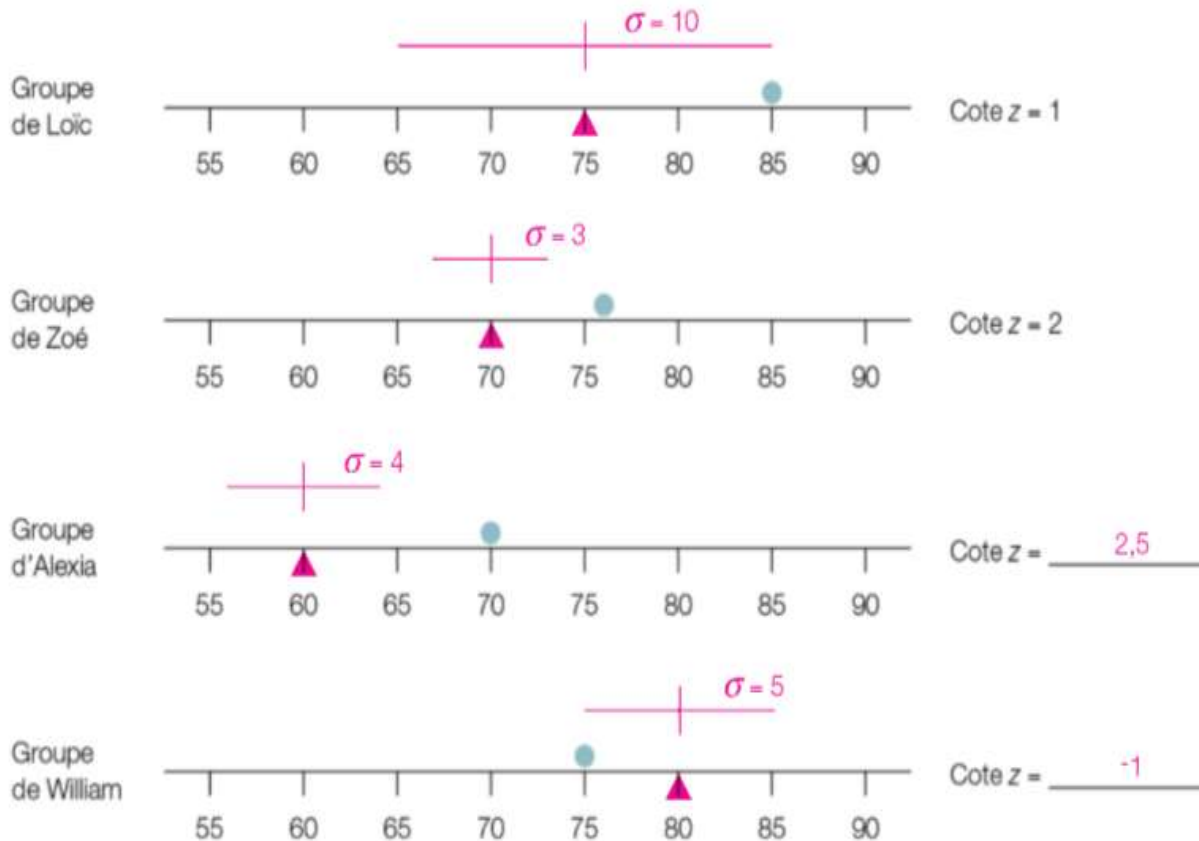
Le CV est exprimé
en pourcentage.

Interprétation

Les coefficients de variation indiquent qu'en 1930, il y avait beaucoup plus de disparité entre les salaires qu'en 2012.

En 2012, la distribution des salaires est homogène, car le $CV < 15\%$.

La cote z : permet de déterminer la position de la note de chaque étudiant par rapport aux autres notes du groupe



Cote z
$\text{Cote } z = \frac{\text{valeur} - \text{moyenne}}{\text{écart type}}$

Une cote z de 2 ou plus à un examen est considérée comme une bonne note

Calculer CV et cote z en se basant sur exemple p. 83

**Répartition des propriétaires
d'une PME selon leur âge au démarrage
de l'entreprise, Québec, 2013**

Âge au démarrage de l'entreprise	Pourcentage de propriétaires
Moins de 25 ans	16,4 %
[25 ans; 35 ans[35,1 %
[35 ans; 45 ans[28,8 %
[45 ans; 55 ans[16,2 %
55 ans et plus	3,5 %
Total	100,0 %

Source: Fondation de l'entrepreneursip. *Indice entrepreneurial québécois 2013. Les entrepreneurs du Québec font-ils preuve d'audace ?*

	A	B	C	D	E	F	G	H
	Répartition des propriétaires d'une PME selon leur âge au démarrage de l'entreprise, Québec, 2013							
	âge	Pourcentage (fi)	ci	cifi	(ci-μ)	(ci-μ) ²	(ci-μ)2fi	
5	Moins de 25	16,40%	20	3,28	-15,5	240,25	39,401	
7	[25;35[35,10%	30	10,53	-5,5	30,25	10,61775	
8	[35;45[28,80%	40	11,52	4,5	20,25	5,832	
9	[45;55[16,20%	50	8,1	14,5	210,25	34,0605	
0	55 ans et plu	3,50%	60	2,1	24,5	600,25	21,00875	
1				35,53			110,92	
2			valeur arrondie de la moyenne	35,5		écart type	10,5318564	
3						CV	29,7%	
4	a)							
5								
6								
7		i) $\mu = \sum cifi$						
8		$\mu = 35,5$ ans						
9		$\sigma = 10,5$ ans						
0								
1								
2								

Interprétation:

On peut estimer que les propriétaires d'une PME avaient en moyenne 35,5 ans.

ii) Pour savoir si la distribution est homogène ou pas, il faut calculer le coefficient de variation CV

$$CV = \frac{\sigma}{\mu} \times 100\%$$

$$CV = 29,7\%$$

$$\text{Cote } z = \frac{58 - 35,5}{10,5} = 2,1$$

L'âge de cette personne a une cote z de 2,1 écart type. Il est assez rare de lancer une entreprise à l'âge de 58 ans.

Mesures de position : les quantiles

Les **quantiles** sont des valeurs qui partagent une distribution en un certain nombre de parties égales. Les plus utilisés sont :

- les **quartiles** (Q_1, Q_2, Q_3), qui partagent une distribution en quatre parties comprenant 25 % des données ;
- les **quintiles** (V_1, V_2, V_3, V_4), qui partagent une distribution en cinq parties comprenant 20 % des données ;
- les **déciles** (D_1, D_2, \dots, D_9), qui partagent une distribution en dix parties comprenant 10 % des données ;
- les **centiles** (C_1, C_2, \dots, C_{99}), qui partagent une distribution en cent parties comprenant 1 % des données.

Calcul des quartiles: Q1, Q2 et Q3

- Le premier quartile Q1 est déterminé tel que 25% des données sont situées avant Q1 (inférieur ou égal à Q1)
- Le deuxième quartile Q2= Me
- Le troisième quartile Q3 est déterminé tel que 75% des données sont situées avant Q3 (inférieur ou égal à Q3)

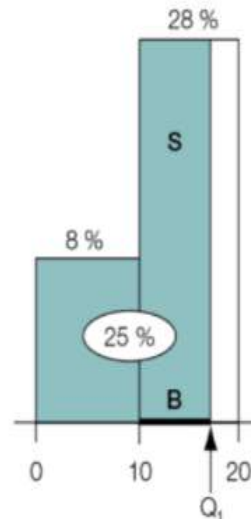
Remarque : le calcul des quartiles se fait de la même manière que la médiane. Il faut procéder par le calcul du **pourcentage cumulé croissant**.

Exemple de calcul de Q1 (manuel, p. 56)

Répartition des arbres selon l'âge

Âge (en ans)	Pourcentage
[0; 10[8 %
[10; 20[28 %
[20; 30[32 %
[30; 40[20 %
[40; 50[12 %
Total	100 %

Esquisse de l'histogramme



- Surface: $25 \% = 8 \% + S$

$$S = 25 \% - 8 \%$$

$$S = 17 \%$$

- $Q_1: 10 + B$

On a: $\frac{\text{Surface}}{\text{Base}}$

28 %	→	10 ans
17 %	→	B ans

$$\text{D'où } B = \frac{17 \times 10}{28} = 6,1 \text{ ans}$$

$$Q_1 = 10 + 6,1 = 16,1 \text{ ans}$$

Q1 appartient à la classe [10;20] (car on atteint les 25% des observations)

Si $Q_1 = 16,1$ ans

On estime que 25% des arbres ont moins de 16,1 ans.

EXERCICE 6, p. 59: Calculer Q1 et Q3.

Répartition des bénéficiaires du Programme de prêts et bourses de la formation technique selon le montant¹ d'aide attribué, 2011-2012

Montant attribué	Pourcentage
Moins de 2 000 \$	26,3 %
[2 000 \$; 4 000 \$[20,1 %
[4 000 \$; 6 000 \$[17,7 %
[6 000 \$; 8 000 \$[16,3 %
[8 000 \$; 10 000 \$[9,7 %
10 000 \$ et plus	9,9 %
Total	100,0 %

1. Prêt + bourse.

Source: Ministère de l'Enseignement supérieur. *Statistiques. Rapport 2011-2012*, 2014.

Calcul de Q1 et Q3 exercice 6 p. 59.

Répartition des bénéficiaires du Programme de prêts et bourses de la formation technique selon le montant		
Montant attribué	Pourcentage	Pourcentage cumulé
[0; 2000]	26,30%	26,3%
[2000; 4000]	20,10%	46,4%
[4000; 6000]	17,70%	64,1%
[6000; 8000]	16,30%	80,4%
[8000; 10 000]	9,70%	90,1%
[10 000; 12 000]	9,90%	100,0%

Détermination du 1er Quartile:

Q1 appartient à la classe [0; 2000] (car on a atteint le pourcentage cumulé de 25%)

26,3%.....2000
25%.....X

$$x = \frac{25\% \times 2000}{26,3\%} = 1901\$$$

Interprétation:

On peut estimer que 25% des bénéficiaires de la formation technique ont reçu moins de 1901\$ du Programme des prêts et bourses en 2011-2012.

Détermination du 3e Quartile:

Q3 appartient à la classe [6000; 8000] (car on a atteint le pourcentage cumulé de 75%)

$$26,3\% + 20,1\% + 17,7\% + S = 75\%$$

$$S = 10,9\%$$

$$\begin{array}{l} 16,3\% \dots\dots\dots 2000 \\ 10,9\% \dots\dots\dots X \end{array}$$

$$x = \frac{10,9\% \times 2000}{16,3\%} = 1337,4\$$$

$$Q3 = 6000 + 1337,4 = 7337,4\$$$

Interprétation:

On peut estimer que 75% des bénéficiaires de la formation technique ont reçu moins de 7337,4\$ du Programme des prêts et bourses en 2011-2012.

La méthode de l'interpolation linéaire permet de retrouver le même résultat. Il s'agit de résoudre l'égalité suivante:

Travail à faire

- Exercices 1.5: 5, 9 et 10, p. 79-80.