

SÉRIE N° 3 : LES CARACTÉRISTIQUES DE DISPERSION
- CORRECTION -

L'étendue - L'écart absolu moyen - La variance - l'écart type - Les intervalles interquantiles - Le coefficient de variation

EXERCICE I :

Groupe A :	8	9	10	11	12
Groupe B :	2.4	4.4	11.4	15.4	16.4

1. Calcul de la moyenne pour chaque groupe

La **moyenne arithmétique** est donnée par la formule :

$$\bar{x} = \frac{\sum x_i}{n}$$

Groupe A :

Nombre de valeurs : n=5

$$\bar{x}_A = \frac{8 + 9 + 10 + 11 + 12}{5} = \frac{50}{5} = 10$$

Groupe B :

Nombre de valeurs : n=5

$$\bar{x}_B = \frac{2.4 + 4.4 + 11.4 + 15.4 + 16.4}{5} = \frac{50}{5} = 10$$

Les moyennes des deux groupes sont identiques : 10

2. Lequel des deux groupes est plus homogène ?

L'homogénéité est évaluée par une **mesure de dispersion**, comme l'**écart-type**, qui mesure à quel point les valeurs sont éloignées de la moyenne.

L'écart-type est donné par la formule :

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

- Calcul de l'écart-type pour le **groupe A** :

$$\sigma_A = \sqrt{\frac{(8 - 10)^2 + (9 - 10)^2 + (10 - 10)^2 + (11 - 10)^2 + (12 - 10)^2}{5}}$$

Donc l'écart-type pour le groupe A est **1.41**

- Calcul de l'écart-type pour le **groupe B** :

$$\sigma_B = \sqrt{\frac{(2.4 - 10)^2 + (4.4 - 10)^2 + (11.4 - 10)^2 + (15.4 - 10)^2 + (16.4 - 10)^2}{5}}$$

Donc l'écart-type pour le groupe B est **5.68**

Conclusion : Quel groupe est plus homogène ?

- Le groupe A a un écart-type de **1.41**, donc les notes sont **proches de la moyenne**.
 - Le groupe B a un écart-type de **5.68**, ce qui signifie une **plus grande dispersion** des notes.
- Le groupe A est **plus homogène** car les notes sont plus regroupées autour de la moyenne.
Le groupe B est **plus dispersé**, avec des écarts plus grands entre les notes.

EXERCICE 2:

Groupe A :		Groupe B :	
Notes (Xi)	Effectif (ni)	Notes (Xi)	Effectif (ni)
8	2	6	2
9	2	8	2
10	1	9	2
11	1	13	1
		14	1

La **moyenne pondérée** est donnée par la formule :

$$\bar{x} = \frac{\sum (x_i \times n_i)}{\sum n_i}$$

où :

- x_i est la valeur des notes,
- n_i est l'effectif correspondant.

Groupe A :

- **Total des effectifs :**

$$\sum n_i = 2 + 2 + 1 + 1 = 6$$

- **Moyenne :**

$$\begin{aligned}\bar{x}_A &= \frac{\sum (x_i \times n_i)}{\sum n_i} = \frac{(8 \times 2) + (9 \times 2) + (10 \times 1) + (11 \times 1)}{6} \\ &= \frac{16 + 18 + 10 + 11}{6} = \frac{55}{6} \approx 9.17\end{aligned}$$

- **L'écart-type :**

$$\sigma = \sqrt{\frac{\sum n_i (x_i - \bar{x})^2}{\sum n_i}}$$

Substituons les valeurs, on trouve : **L'écart-type = 1.07**

Groupe B :

- **Total des effectifs :**

$$\sum n_i = 2 + 2 + 2 + 1 + 1 = 8$$

- **Moyenne :**

$$\bar{x}_B = \frac{\sum (x_i \times n_i)}{\sum n_i} = \frac{(6 \times 2) + (8 \times 2) + (9 \times 2) + (13 \times 1) + (14 \times 1)}{8}$$

$$= \frac{12 + 16 + 18 + 13 + 14}{8} = \frac{73}{8} \approx 9.13$$

- **L'écart-type :**

$$\sigma = \sqrt{\frac{\sum n_i (x_i - \bar{x})^2}{\sum n_i}}$$

Substituons les valeurs, on trouve : **L'écart-type = 2.76**

Comparaison des deux groupes

- Les moyennes sont **presque identiques** :
 - Groupe A : **≈ 9.17**
 - Groupe B : **≈ 9.13**
- **L'écart-type du groupe B est plus grand** :
 - Groupe A : **≈ 1.07**
 - Groupe B : **≈ 2.76**

Le groupe A est plus homogène car son écart-type est plus petit. Les notes sont plus concentrées autour de la moyenne.

Le groupe B est plus dispersé, avec des notes qui varient plus largement autour de la moyenne.

EXERCICE 3:

	Salaire moyen (DH)	Variance	Pourcentage
Hommes	6000	1200	0.6
Femmes	3000	800	0.4

1. Calcul du salaire moyen dans l'entreprise :

Le salaire moyen total est donné par la moyenne pondérée des salaires moyens des hommes et des femmes :

$$\bar{X} = p_H \cdot \bar{X}_H + p_F \cdot \bar{X}_F$$

- $p_H = 0.6$ (proportion des hommes),
- $p_F = 0.4$ (proportion des femmes),
- $\bar{X}_H = 6000$ DH (salaire moyen des hommes),
- $\bar{X}_F = 3000$ DH (salaire moyen des femmes).

$$\bar{X} = (0.6 \times 6000) + (0.4 \times 3000)$$

$$\bar{X} = 3600 + 1200 = 4800 \text{ DH}$$

Le **salaire moyen** dans l'entreprise est de **4800 DH**.

2. Comparaison de la dispersion des salaires :

Puisque nous avons **deux moyennes différentes**, il est plus pertinent d'utiliser le **coefficient de variation (CV)** plutôt que l'écart-type.

Le **coefficient de variation (CV)** permet de comparer la dispersion **en termes relatifs** en tenant compte du salaire moyen. Il est défini par :

$$CV = \frac{\sigma}{\bar{X}} \times 100$$

Calcul du CV pour les hommes :

$$\sigma_H = \sqrt{\text{Var}_H} = \sqrt{1200} \approx 34.64$$

$$CV_H = \frac{34.64}{6000} \times 100 \approx 0.58\%$$

Calcul du CV pour les femmes :

$$\sigma_F = \sqrt{\text{Var}_F} = \sqrt{800} \approx 28.28$$

$$CV_F = \frac{28.28}{3000} \times 100 \approx 0.94\%$$

Le coefficient de variation (CV) est **plus élevé pour les femmes (0.94% contre 0.58% pour les hommes)**, ce qui signifie que **la dispersion relative des salaires est plus importante chez les femmes**.

3. Calcul de la variance totale des salaires :

La variance totale est donnée par la formule :

$$\sigma_T^2 = P_H \times (\sigma_H^2 + (\bar{X}_H - \bar{X})^2) + P_F \times (\sigma_F^2 + (\bar{X}_F - \bar{X})^2)$$

où :

- $\sigma_H^2 = 1200, \sigma_F^2 = 800,$
- $\bar{X}_H - \bar{X} = 6000 - 4800 = 1200,$
- $\bar{X}_F - \bar{X} = 3000 - 4800 = -1800.$

Substituons les valeurs, on trouve :

La variance totale des salaires dans l'entreprise est 2 161 040 DH²

EXERCICE 4:

	Plus petit salaire	Q1	Médiane	Q3	Plus grand salaire
Entreprise A	1300	1600	1800	2400	6500
Entreprise B	1350	1400	1800	3100	6800

1. Calcul de l'étendue des salaires

L'**étendue** est donnée par la différence entre le plus grand et le plus petit salaire :

- Entreprise A :
6500-1300=5200
- Entreprise B :
6800-1350=5450

L'étendue des salaires est de **5200€** pour l'entreprise A et **5450€** pour l'entreprise B.

2. Calcul de l'intervalle interquartile (IQ)

L'intervalle interquartile (IQ) mesure la dispersion des salaires entre le 1er quartile (Q1) et le 3e quartile(Q3) :

$$\text{IQ} = \text{Q3} - \text{Q1}$$

- Entreprise A :
2400-1600=**800**
- Entreprise B :
3100-1400=**1700**

L'intervalle interquartile est de 800€ pour l'entreprise A et 1700€ pour l'entreprise B.

3. Comparaison des distributions des salaires

- L'entreprise B a une étendue légèrement plus grande (5450€ contre 5200€), ce qui indique que les écarts entre les salaires extrêmes y sont plus marqués.
- L'entreprise B a un IQ beaucoup plus large (1700€ contre 800€), ce qui signifie que les salaires intermédiaires y sont plus dispersés.
- L'entreprise A présente un IQ plus faible, ce qui suggère que la majorité des salaires sont plus concentrés autour de la médiane.

Quelle entreprise offre une distribution plus équilibrée ?

- L'entreprise A a une distribution plus équilibrée, car son IQ est plus faible, ce qui indique que la majorité des employés ont des salaires moins dispersés autour de la médiane.

EXERCICE 5 :

Salaires	Nbr salariés
[2000-3000[	10
[3000-5000[	100
[5000-6000[	60
[6000-8000[	20
[8000-10000[	10

Total : 200

1. Étendue des salaires

L'étendue est la différence entre le salaire maximum et le salaire minimum :

$$E = S_{\max} - S_{\min} = 10000 - 2000 = 8000$$

Interprétation : L'étendue montre l'amplitude totale des salaires dans l'échantillon. Une étendue élevée indique une grande dispersion entre les salaires les plus bas et les plus hauts.

2. Moyenne des salaires

La moyenne est calculée en utilisant les classes de salaires et leurs effectifs. On prend comme valeur représentative de chaque classe le centre de l'intervalle :

Salaires	Nbr salariés (ni)	Centre ci	Amplitude ai	ni x ci	Ni cc	fi	Ficc
[2000-3000[	10	2500	1000	25000	10	0.05	0.05
[3000-5000[	100	4000	2000	400000	110	0.5	0.55
[5000-6000[	60	5500	1000	330000	170	0.3	0.85
[6000-8000[	20	7000	2000	140000	190	0.1	0.95
[8000-10000[	10	9000	2000	90000	200	0.05	1
N	200						

$$\bar{X} = \frac{\sum n_i \times c_i}{N}$$

Substituons les valeurs, on trouve : **Le salaire moyen est de 4925 DH**

Interprétation : Le salaire moyen dans cette distribution est de **4925**, ce qui donne une idée centrale du niveau des salaires.

3. Premier et troisième quartile (Q1 et Q3)

Les quartiles sont les valeurs qui divisent la distribution en quatre parties égales.

Le premier quartile correspond à la valeur pour laquelle 25 % des salariés ont un salaire inférieur ou égal à cette valeur.

- **Effectif total** : 200.
- **Position de Q1** : $0,25 \times 200 = 50$.
- L'effectif cumulé nous indique que le **50^e salarié est dans la classe [3000-5000[**.

$$Q1 = b_{i-1} + a_i \left(\frac{n/4 - N_{i-1}^+}{n_i} \right)$$

Substituons les valeurs, on trouve : **Q1=3800**

Le troisième quartile correspond à la valeur pour laquelle 75 % des salariés ont un salaire inférieur ou égal à cette valeur.

- **Position de Q3** : $0,75 \times 200 = 150$.
On cherche la classe où le 150^e salarié se situe. L'effectif cumulé nous indique que le **150^e salarié est dans la classe [5000-6000[**.

$$Q3 = b_{i-1} + a_i \left(\frac{3n/4 - N_{i-1}^+}{n_i} \right)$$

Substituons les valeurs, on trouve : **Q3= 5666,67**

4. L'intervalle interquartile (IQ) est la différence entre Q3 et Q1.

$$IQ = Q3 - Q1 = 5666,67 - 3800 = 1866,67$$

Rôle : L'intervalle interquartile mesure la dispersion des valeurs centrales (entre Q1 et Q3). Il permet d'ignorer les valeurs extrêmes et de mieux comprendre la concentration des salaires autour de la médiane. Ici, **50% des salariés gagnent un salaire compris entre 3800 DH et 5666.67 DH.**

5. Calculer l'écart-type des salaires et l'interpréter

L'écart-type mesure la dispersion des salaires autour de la moyenne.

Formule :

$$\sigma = \sqrt{\frac{\sum n_i (c_i - \bar{X})^2}{N}}$$

Substituons les valeurs, on trouve : **L'écart-type est 1442.87 DH**

Donc, **l'écart-type est 1442.87 DH**. Cela signifie qu'en moyenne, les salaires s'écartent de **1442.87 DH** par rapport à la moyenne.

6. Calcul du coefficient de variation (CV)

Le coefficient de variation est donné par :

$$CV = \frac{\sigma}{\bar{X}} \times 100$$

$$CV = \frac{1442.87}{4925} \times 100$$

$$CV \approx 29.30\%$$

Un **CV de 29.30%** indique une **dispersion modérée** des salaires.

7. Utilisation du coefficient de variation (CV)

Le CV est utilisé pour :

- Comparer la dispersion de plusieurs distributions avec des unités ou des échelles différentes.
- Mesurer la **relative** dispersion par rapport à la moyenne.
- Évaluer la stabilité d'un phénomène : **plus le CV est élevé, plus la variabilité est forte.**

En général :

- ✓ **CV < 15%** → Faible dispersion
- ✓ **15% ≤ CV ≤ 30%** → Dispersion modérée
- ✓ **CV > 30%** → Forte dispersion

Ici, **29.30%** indique une **dispersion relativement élevée des salaires**, mais elle reste **modérée** selon l'échelle d'interprétation.