

Statistique descriptive

Chapitre 8 : Les indicateurs de dispersion



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Pourquoi la moyenne ne suffit pas

Les mesures simples

La variance et l'écart-type

La dispersion relative

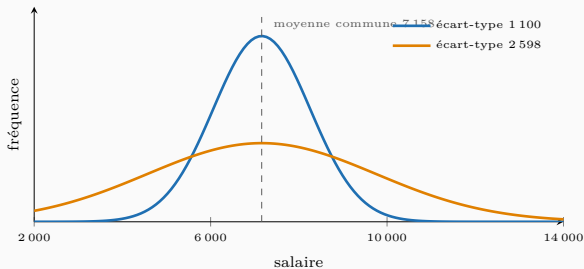
La comparaison des mesures

Ce qu'il faut retenir

Pourquoi la moyenne ne suffit pas

Deux entreprises, une même moyenne

deux entreprises, la même moyenne, deux réalités



la moyenne seule ne dit rien
de ce que gagne un salarié
pris au hasard.

**Une moyenne sans dispersion
n'est pas une information.**

$CV = \frac{\sigma}{\bar{x}}$ permet de comparer
deux dispersions exprimées
dans des unités différentes.

Ce qu'on cherche

Un **indicateur de dispersion** mesure à quel point les observations s'écartent les unes des autres, ou de leur centre.

Il répond à : la moyenne est-elle représentative ?

La conséquence pratique

Deux entreprises annoncent un salaire moyen de **7 158** dirhams. Dans la première, presque tout le monde gagne entre **6 000** et **8 000**. Dans la seconde, les salaires vont de **2 800** à **19 000**.

Ce ne sont pas les mêmes entreprises, et un candidat n'y court pas le même risque.

Une moyenne publiée sans mesure de dispersion n'est pas une information — c'est au mieux un ordre de grandeur.

Les mesures simples

Étendue

$$E = x_{\max} - x_{\min}$$

Sur nos salaires

$$E = 19\,050 - 2\,810 = 16\,240 \text{ dirhams}$$

Simple, et presque inutilisable

Elle ne dépend que de **deux** observations sur 500 — les deux plus extrêmes, c'est-à-dire les moins représentatives.

Pire : elle **croît mécaniquement** avec n . Plus on observe d'individus, plus on a de chances de rencontrer une valeur extrême. Comparer les étendues de deux échantillons de tailles différentes n'a donc aucun sens.

On la mentionne parce qu'elle est intuitive, jamais pour conclure.

Écart absolu moyen

$$e = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|$$

Sur nos données

$$e = 1\,972 \text{ dirhams}$$

Un salarié s'écarte en moyenne de 1 972 dirhams du salaire moyen.

C'est l'indicateur le plus naturel — et il n'est presque jamais employé

Il répond exactement à la question posée, et son interprétation est directe.

Il a pourtant un défaut rédhibitoire : la **valeur absolue ne se dérive pas** et ne se manipule pas algébriquement. Impossible d'en tirer des formules, de la décomposer par sous-groupes, de la faire entrer dans un calcul d'intervalle de confiance.

D'où le recours au **carré**, qui remplit la même fonction — annuler les signes — tout en restant dérivable.

La variance et l'écart-type

Variance

$$V(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

C'est la **moyenne des carrés des écarts à la moyenne**.

Écart-type

$$\sigma_x = \sqrt{V(x)}$$

On prend la racine pour revenir à l'**unité de la variable**.

Sur nos salaires

$$V = 6\,750\,122 \text{ dirhams}^2 \qquad \sigma = 2\,598 \text{ dirhams}$$

Pourquoi l'écart-type et pas la variance

La variance s'exprime en **dirhams au carré**, une unité qui ne veut rien dire. On ne peut ni la comparer à la moyenne, ni la reporter sur un graphique.

L'écart-type est en dirhams : il se compare directement à la moyenne, se lit sur l'axe, s'ajoute et se retranche.

On calcule avec la variance, on communique avec l'écart-type.

Théorème de König-Huygens

$$V(x) = \frac{1}{n} \sum x_i^2 - \bar{x}^2 = \overline{x^2} - \bar{x}^2$$

La moyenne des carrés moins le carré de la moyenne.

La démonstration, en trois lignes

$$\begin{aligned} \frac{1}{n} \sum (x_i - \bar{x})^2 &= \frac{1}{n} \sum (x_i^2 - 2\bar{x}x_i + \bar{x}^2) \\ &= \frac{1}{n} \sum x_i^2 - 2\bar{x} \underbrace{\frac{1}{n} \sum x_i}_{=\bar{x}} + \bar{x}^2 = \overline{x^2} - 2\bar{x}^2 + \bar{x}^2 = \overline{x^2} - \bar{x}^2 \end{aligned}$$

Vérification numérique

$$\overline{x^2} = 57\,990\,808 \quad \bar{x}^2 = 7\,158,26^2 = 51\,240\,686$$

$$V = 57\,990\,808 - 51\,240\,686 = 6\,750\,122 \quad \checkmark$$

Pourquoi cette formule compte, au-delà du calcul

Elle ne demande qu'un **seul passage** sur les données : on cumule $\sum x_i$ et $\sum x_i^2$ en même temps. La définition, elle, exige de connaître $\bar{x}$ avant de commencer.

C'est la formule qu'utilisent les tableurs et les calculatrices — et surtout, c'est celle dont **la covariance sera la généralisation** au chapitre 13. Retenez sa structure : *moyenne du produit moins produit des moyennes*.

Trois propriétés à connaître

1. $V(x) \geq 0$, et $V(x) = 0$ si et seulement si toutes les valeurs sont égales.
2. $V(x + b) = V(x)$: **ajouter une constante ne change pas la dispersion.**
3. $V(ax) = a^2 V(x)$, donc $\sigma_{ax} = |a| \sigma_x$.

La propriété 2 en clair

Une augmentation générale de 500 dirhams pour tous **ne réduit pas les inégalités** : l'écart-type est inchangé.

Une augmentation de 10 % pour tous multiplie l'écart-type par 1,1 : elle **augmente** l'écart absolu entre les salariés.

Ces deux faits, qui reviennent à chaque négociation salariale, se lisent directement dans les propriétés de la variance.

La dispersion relative

Coefficient de variation

$$CV = \frac{\sigma_x}{\bar{x}}$$

C'est un nombre **sans unité**, souvent exprimé en pourcentage. Il exige $\bar{x} > 0$.

Comparer deux variables incomparables

Variable	Moyenne	Écart-type	<i>CV</i>
Salaire	7 158 dh	2 598 dh	0,363
Ancienneté	10,14 ans	5,96 ans	0,588

Les écarts-types, l'un en dirhams et l'autre en années, ne se comparent pas. Les coefficients de variation, si.

L'ancienneté est nettement plus dispersée que le salaire, relativement à son niveau.

Ses deux usages en gestion

1. **Comparer des variables d'unités différentes**, comme ci-dessus.
2. **Comparer des populations de niveaux très différents** : la dispersion des salaires d'une multinationale et d'une PME.

En finance, le CV porte un autre nom et s'inverse : le rapport de la rentabilité moyenne à sa volatilité est le **ratio de Sharpe**. C'est la même idée — rapporter la dispersion au niveau — au service d'une autre question.

Ordres de grandeur

CV	Lecture usuelle
$< 0,15$	population très homogène
$0,15 \text{ à } 0,30$	dispersion modérée
$0,30 \text{ à } 0,60$	dispersion forte
$> 0,60$	population très hétérogène

Ces seuils sont indicatifs, pas normatifs

Ils dépendent entièrement du domaine. Un CV de 0,05 serait énorme pour un processus industriel sous contrôle, et dérisoire pour des rentabilités boursières.

Le bon repère est toujours la comparaison : le même indicateur, sur une autre population ou une autre période.

La comparaison des mesures

Le tableau récapitulatif

Les cinq indicateurs de dispersion

Indicateur	Formule	Robuste ?	Unité
Étendue	$x_{\max} - x_{\min}$	non	celle de x
Intervalle interquartile	$Q_3 - Q_1$	oui	celle de x
Écart absolu moyen	$\frac{1}{n} \sum x_i - \bar{x} $	assez	celle de x
Variance	$\frac{x^2}{n} - \bar{x}^2$	non	celle de x au carré
Écart-type	$\sqrt{V}$	non	celle de x
Coefficient de variation	$\sigma / \bar{x}$	non	aucune

Toutes les valeurs, sur nos salaires

Indicateur	Valeur
Étendue	16 240
<i>IQ</i>	3 012
Écart absolu moyen	1 972
Écart-type	2 598
<i>CV</i>	0,363

Deux remarques sur ces nombres

L'écart-type dépasse l'écart absolu moyen : 2 598 contre 1 972. C'est toujours le cas, parce que le carré donne plus de poids aux grands écarts.

L'écart entre les deux est d'ailleurs un indice d'asymétrie : plus la distribution a une queue longue, plus l'écart-type s'éloigne de l'écart absolu moyen.

En pratique

- Pour **décrire** une distribution asymétrique ou comportant des valeurs extrêmes : IQ et boîte à moustaches.
- Pour **calculer** quoi que ce soit ensuite : variance et écart-type. Ils sont les seuls à entrer dans les formules de la covariance, de la corrélation et de toute l'inférentielle.
- Pour **comparer** des unités différentes : coefficient de variation.

Et dans un tableau de résultats professionnel, on donne les trois.

Ce qu'il faut retenir

Six points

1. Une moyenne sans dispersion n'est pas une information.
2. L'**étendue** ne repose que sur deux observations et croît avec n : à éviter.
3. **Variance** : moyenne des carrés des écarts. **Écart-type** : sa racine, dans l'unité de la variable.
4. **König-Huygens** : $V(x) = \overline{x^2} - \bar{x}^2$ — moyenne du produit moins produit des moyennes.
5. $V(x + b) = V(x)$ et $V(ax) = a^2 V(x)$: une prime uniforme ne réduit pas les écarts.
6. **Coefficient de variation** $CV = \sigma / \bar{x}$: sans unité, il permet de comparer l'incomparable.

La suite

Deux distributions peuvent avoir la même moyenne **et** le même écart-type sans se ressembler : l'une symétrique, l'autre étalée d'un côté.

Il faut donc aller plus loin — au moment d'ordre trois, puis d'ordre quatre. C'est le chapitre 9.