

Statistique descriptive

Chapitre 9 : Les indicateurs de forme



Jaouad Madkour

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

Au-delà du centre et de la dispersion

L'asymétrie

L'aplatissement

En pratique

Ce qu'il faut retenir

Au-delà du centre et de la
dispersion

Deux distributions, même moyenne, même écart-type

Et pourtant deux réalités différentes : l'une symétrique, l'autre étalée d'un côté; l'une aux extrêmes rares, l'autre aux extrêmes fréquents.

Il faut donc deux indicateurs supplémentaires, qui répondent à deux questions distinctes :

- l'**asymétrie** : la distribution penche-t-elle d'un côté ?
- l'**aplatissement** : les valeurs extrêmes sont-elles fréquentes ?

Moment centré d'ordre r

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

Les quatre premiers

Ordre	Valeur	Ce qu'il mesure
$r = 1$	$m_1 = 0$	rien — c'est la propriété d'équilibre
$r = 2$	$m_2 = V(x)$	la dispersion
$r = 3$	m_3	l' asymétrie
$r = 4$	m_4	l' aplatissement

Pourquoi les ordres impairs mesurent l'asymétrie

Une puissance **impaire** conserve le signe : un écart de -3 contribue pour -27 , un écart de $+3$ pour $+27$. Ils se compensent si la distribution est symétrique, et ne se compensent pas sinon.

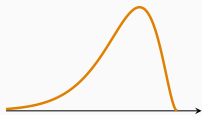
Une puissance **paire** détruit le signe : $(-3)^4$ et $(+3)^4$ valent tous deux 81. Elle ne peut donc rien dire du sens, seulement de l'ampleur.

C'est toute la logique des moments : l'ordre trois détecte le déséquilibre, l'ordre quatre l'épaisseur des queues.

L'asymétrie

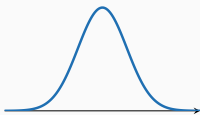
asymétrie **négative**

$\gamma_1 < 0$: étalée à gauche



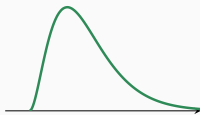
symétrique

$\gamma_1 = 0$: moyenne = médiane



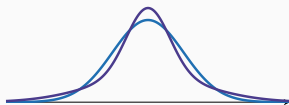
asymétrie **positive**

$\gamma_1 > 0$: étalée à droite



l'aplatissement

— normale, $\gamma_2 = 3$
— leptokurtique, $\gamma_2 > 3$



Deux questions distinctes, deux moments.

L'*asymétrie* demande : la distribution penche-t-elle d'un côté ? Elle se mesure par le moment centré d'ordre **trois**, réduit : $\gamma_1 = m_3/\sigma^3$.

L'*aplatissement* demande : les valeurs extrêmes sont-elles plus fréquentes que sous une loi normale ? Moment d'ordre **quatre** : $\gamma_2 = m_4/\sigma^4$.

Sur nos salaires : $\gamma_1 = 1,16$ et $\gamma_2 = 4,84$. Étalée à droite, et des extrêmes plus fréquents que sous la normale.

Coefficient d'asymétrie

$$\gamma_1 = \frac{m_3}{\sigma^3}$$

On divise par σ^3 pour obtenir un nombre **sans unité**, comparable d'une variable à l'autre.

La lecture

Valeur	Forme	Position des indicateurs
$\gamma_1 = 0$	symétrique	mode = médiane = moyenne
$\gamma_1 > 0$	étalée à droite	mode < médiane < moyenne
$\gamma_1 < 0$	étalée à gauche	mode > médiane > moyenne

Sur nos salaires

$$\gamma_1 = 1,16$$

Nettement positif : distribution étalée à droite, ce que confirment $\text{mode } 5\,750 < \text{médiane } 6\,770 < \text{moyenne } 7\,158$.

Un ordre de grandeur

En pratique, on considère que la distribution est **à peu près symétrique** si $|\gamma_1| < 0,5$, **modérément asymétrique** entre 0,5 et 1, et **fortement asymétrique** au-delà de 1.

Notre valeur de 1,16 range donc les salaires dans la dernière catégorie — ce qui est la règle générale pour tous les revenus.

Deux mesures alternatives

Coefficient de Pearson — fondé sur l'écart entre moyenne et mode :

$$A_P = \frac{\bar{x} - \text{mode}}{\sigma}$$

Coefficient de Yule — fondé sur les quartiles :

$$A_Y = \frac{(Q_3 - Me) - (Me - Q_1)}{Q_3 - Q_1} = \frac{Q_3 + Q_1 - 2Me}{Q_3 - Q_1}$$

Les trois, sur nos données

Coefficient	Valeur
Fisher γ_1	1,16
Pearson A_P	0,54
Yule A_Y	0,036

Les autres coefficients (suite 2)

Pourquoi ils diffèrent autant

Ils ne mesurent pas la même chose.

Yule ne regarde que la boîte : la médiane est presque au centre de l'intervalle interquartile, d'où une valeur quasi nulle. Le cœur de la distribution est presque symétrique.

Fisher utilise toutes les valeurs, en donnant un poids au cube aux plus éloignées : il capte pleinement la queue droite.

La comparaison des deux est en elle-même un diagnostic : $A_Y \simeq 0$ et $\gamma_1 = 1,16$ signifient que l'asymétrie ne vient pas du centre de la distribution mais **uniquement de sa queue**. Autrement dit : la masse des salariés est répartie symétriquement, et quelques hauts salaires créent tout l'étalement.

Ce que cela vaut pour une négociation

« Les salaires sont-ils inégalement répartis ? » — la réponse dépend de qui l'on regarde.

Pour la majorité du personnel, la répartition est équilibrée ($A_Y \simeq 0$). Ce sont les rémunérations les plus élevées qui créent l'asymétrie ($\gamma_1 > 1$).

Un seul indicateur aurait manqué cette nuance. C'est un argument pour toujours en calculer deux.

L'aplatissement

Coefficient d'aplatissement

$$\gamma_2 = \frac{m_4}{\sigma^4}$$

La référence est la **loi normale**, pour laquelle $\gamma_2 = 3$.

La lecture		
Valeur	Nom	Interprétation
$\gamma_2 = 3$	mésokurtique	comme la loi normale
$\gamma_2 > 3$	leptokurtique	pic plus aigu, queues épaisses
$\gamma_2 < 3$	platykurtique	distribution plus plate

La notation à surveiller

Certains logiciels affichent l'**excès d'aplatissement**, $\gamma_2 - 3$, dont la référence est donc 0 et non 3.

R (**kurtosis** de **moments**) donne γ_2 . Python (**scipy.stats.kurtosis**) donne l'excès par défaut, sauf si l'on précise **fisher=False**.

Vérifiez toujours la convention avant d'interpréter. Une valeur de 1,84 peut signifier « moins aplatie que la normale » ou « nettement plus » selon le logiciel.

Sur nos salaires

$$\gamma_2 = 4,84 \quad \text{soit un excès de } 1,84$$

Distribution **leptokurtique** : les salaires très éloignés de la moyenne sont plus fréquents que sous une loi normale de même écart-type.

Ce que l'aplatissement dit vraiment

Une idée fausse répandue

On enseigne souvent que γ_2 mesure « la hauteur du pic ». C'est inexact.

Le moment d'ordre quatre donne un poids en $(x - \bar{x})^4$: les valeurs proches de la moyenne, dont l'écart est petit, contribuent **presque rien**. Ce sont les valeurs éloignées qui font toute la valeur du coefficient.

L'aplatissement mesure l'épaisseur des queues, pas la forme du sommet.

Pourquoi c'est l'indicateur central de la finance

Les rentabilités boursières ont un γ_2 compris entre 5 et 10, parfois davantage. Les krachs sont donc bien plus fréquents que ne le prédirait une loi normale.

Tous les modèles de risque de marché qui supposent la normalité **sous-estiment systématiquement** la probabilité des pertes extrêmes. C'est l'un des reproches faits à la Value at Risk gaussienne, et l'une des raisons de la crise de 2008.

Ce simple coefficient, calculé sur une série de rentabilités, est donc le premier test à faire avant tout modèle de risque.

En pratique

Les cinq nombres à donner

Question	Indicateur	Valeur
Où ?	moyenne	7 158
Où ? (robuste)	médiane	6 770
Combien dispersée ?	écart-type	2 598
Dispersion relative	CV	0,363
Penche-t-elle ?	γ_1	1,16
Queues épaisses ?	γ_2	4,84

Ce que ces six nombres racontent

Une distribution de salaires centrée autour de 6 800 dirhams, assez dispersée, **nettement étalée vers les hauts salaires**, et comportant des valeurs élevées plus fréquentes que sous une loi normale.

Ce diagnostic tient en une phrase, se calcule en une ligne de code — et il est reproductible par n'importe qui à partir des mêmes données. C'est exactement l'objectif de la statistique descriptive.

Les moments sont fragiles

γ_1 élève les écarts au cube, γ_2 à la puissance quatre. Une **seule** valeur aberrante peut donc les bouleverser.

Sur nos données, remplacer le plus haut salaire par 190 050 dirhams ferait passer γ_2 de 4,8 à 414 — une seule ligne modifiée sur cinq cents.

Toujours regarder l'histogramme et la boîte à moustaches avant d'interpréter un coefficient de forme.
Une valeur extrême de γ_2 signale bien plus souvent une erreur de saisie qu'une propriété de la population.

Ce qu'il faut retenir

Six points

1. **Moment centré d'ordre r** : $m_r = \frac{1}{n} \sum (x_i - \bar{x})^r$. Ordre 2 : dispersion. Ordre 3 : asymétrie. Ordre 4 : aplatissement.
2. Ordres **impairs** pour le sens, ordres **pairs** pour l'ampleur.
3. $\gamma_1 = m_3/\sigma^3$: positif, la distribution est étalée à **droite**.
4. $\gamma_2 = m_4/\sigma^4$: la référence est **3**. Au-delà, **queues épaisses**.
5. L'aplatissement mesure l'épaisseur des **queues**, pas la hauteur du pic.
6. Ces coefficients sont **très sensibles aux valeurs extrêmes** : jamais sans le graphique.

La suite

Il reste une question que ni la dispersion ni la forme ne traitent : **comment la masse totale se répartit-elle entre les individus ?**

Ce n'est pas la même question. C'est celle de la **concentration**, et c'est le chapitre 10.