

Statistique descriptive

Chapitre 7 : Les quantiles



Jaouad Madkour

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

Généraliser la médiane

Le calcul

La boîte à moustaches

Les mesures dérivées

Ce qu'il faut retenir

Généraliser la médiane

Pourquoi s'arrêter à deux parts

La médiane coupe la population en deux moitiés. Rien n'oblige à s'arrêter là.

Couper en quatre donne les **quartiles**, en dix les **déciles**, en cent les **centiles**. On obtient une description bien plus fine que trois indicateurs de tendance centrale.

Quantile d'ordre p

Le **quantile d'ordre p** , noté q_p , est la valeur telle qu'une proportion p des individus lui est inférieure :

$$F(q_p) = p$$

C'est donc la lecture de la courbe cumulative **à l'envers** : on entre par les fréquences, on sort par la variable.

Les trois familles

Nom	Nombre de parts	Notation
Quartiles	4	$Q_1, Q_2 = Me, Q_3$
Déciles	10	$D_1, \dots, D_9$
Centiles	100	$C_1, \dots, C_{99}$

Un piège de vocabulaire

Pour couper en quatre, il faut **trois** traits : il y a trois quartiles, neuf déciles, quatre-vingt-dix-neuf centiles.

Et $Q_2 = D_5 = C_{50} = Me$: ce sont quatre noms de la même valeur.

Le calcul

La méthode

On ordonne les valeurs, puis on cherche celle de rang $p(n + 1)$, en interpolant si ce rang n'est pas entier.

Les quantiles de nos 500 salaires

Quantile	Valeur (dirhams)	Lecture
D_1	4 348	10 % gagnent moins
Q_1	5 318	25 % gagnent moins
Me	6 770	la moitié gagne moins
Q_3	8 330	75 % gagnent moins
D_9	10 473	90 % gagnent moins

La formule générale

$$q_p = e_{i-1} + a_i \times \frac{p - F_{i-1}}{F_i - F_{i-1}}$$

où $[e_{i-1} ; e_i[$ est la classe dans laquelle F franchit p .

C'est exactement la formule de la médiane du chapitre 6, avec p à la place de 0,5.

Le premier quartile, en détail

$p = 0,25$ est franchi dans $[4\,000 ; 6\,000[$, où F passe de 0,062 à 0,376 :

$$Q_1 = 4\,000 + 2\,000 \times \frac{0,250 - 0,062}{0,376 - 0,062} = 4\,000 + 2\,000 \times 0,599 = 5\,197$$

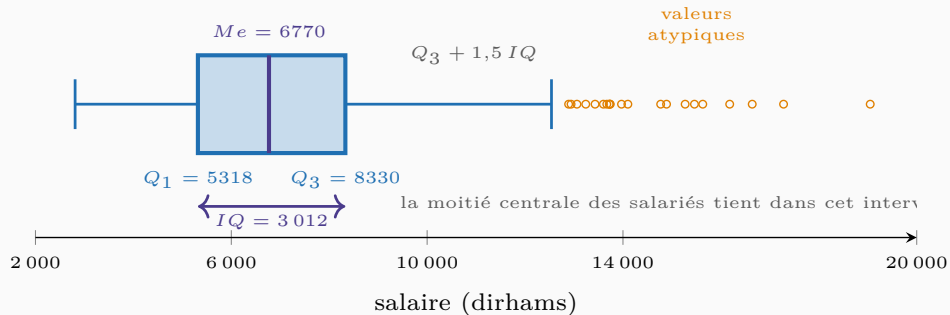
Comparer les deux calculs

	Groupé	Exact	Écart
D_1	4 242	4 348	−2,4 %
Q_1	5 197	5 318	−2,3 %
Me	6 747	6 770	−0,3 %
Q_3	8 532	8 330	+2,4 %
D_9	11 193	10 473	+6,9 %

L'approximation est excellente au centre et se dégrade vers les extrémités — là où les classes sont larges et le remplissage inégal. C'est la limite du travail sur données groupées.

La boîte à moustaches

la boîte à moustaches : cinq nombres résument la distribution



Comment on la lit

- La **boîte** va de Q_1 à Q_3 : elle contient la moitié centrale des individus.
- Le trait intérieur est la **médiane**.
- Les **moustaches** s'étendent jusqu'aux valeurs les plus extrêmes situées à moins de $1,5 IQ$ des quartiles.
- Les points au-delà sont marqués individuellement : ce sont les **valeurs atypiques**.

La règle des 1,5 IQ

$$\text{seuil bas} = Q_1 - 1,5 IQ$$

$$\text{seuil haut} = Q_3 + 1,5 IQ$$

Sur nos salaires

$IQ = 8\,330 - 5\,318 = 3\,012$, donc $1,5\,IQ = 4\,518$.

Seuil haut : $8\,330 + 4\,518 = 12\,848$. Seuil bas : $5\,318 - 4\,518 = 800$, en dessous du minimum observé.

21 salariés dépassent le seuil haut, soit 4,2 % de l'effectif. Aucun en dessous.

Atypique ne veut pas dire faux

Ces 21 salariés ne sont pas des erreurs de saisie : ce sont des cadres expérimentés, parfaitement réels.

La règle des $1,5\,IQ$ est une **convention de repérage**, pas un test. Elle signale des points à examiner, elle ne dit pas qu'ils sont erronés.

Le fait que les valeurs atypiques soient **toutes du même côté** est en soi une information : c'est la signature de l'asymétrie à droite.

Ce qu'une boîte à moustaches montre en un coup d'œil

1. Le **niveau** — la position de la médiane;
2. la **dispersion** — la largeur de la boîte;
3. l'**asymétrie** — la médiane décentrée dans la boîte, une moustache plus longue;
4. les **valeurs atypiques**, individuellement.

Quatre informations dans un dessin de deux centimètres de haut.

Son usage professionnel : comparer

Sa véritable force est la **comparaison**. Placer côte à côte les boîtes des salaires par diplôme, par région, par sexe permet de voir instantanément quels groupes diffèrent en niveau, en dispersion, en asymétrie.

C'est exactement ce que fera le chapitre 12 avec les distributions conditionnelles.

Les mesures dérivées

Trois mesures

Mesure	Formule	Ce qu'elle dit
Intervalle interquartile	$IQ = Q_3 - Q_1$	dispersion de la moitié centrale
Écart interquartile relatif	IQ/Me	dispersion relative , sans unité
Rapport interdécile	D_9/D_1	inégalité entre les extrêmes

Sur nos données

$$IQ = 3\,012 \quad \frac{IQ}{Me} = \frac{3\,012}{6\,770} = 0,445 \quad \frac{D_9}{D_1} = \frac{10\,473}{4\,348} = 2,41$$

Le rapport interdécile est l'indicateur d'inégalité le plus lu

« Les 10 % les mieux payés gagnent 2,41 fois plus que les 10 % les moins payés. »

Cette phrase est immédiatement compréhensible par un non-spécialiste, et c'est pour cela que ce rapport figure dans tous les rapports sociaux d'entreprise et dans les publications sur les inégalités.

À titre de repère, le rapport interdécile des salaires est de l'ordre de 3 en France et dépasse 5 dans les pays très inégalitaires. Notre entreprise, à 2,41, est donc plutôt homogène.

Leur avantage décisif

Toutes ces mesures sont **robustes** : elles ne dépendent que des rangs.

Ajouter un dirigeant à 200 000 dirhams ne change ni Q_1 , ni Q_3 , ni D_9 . L'écart-type, lui, doublerait presque.

C'est pourquoi les organismes statistiques privilégient les quantiles pour décrire les revenus.

Les limites

1. Elles **ignorent** ce qui se passe à l'intérieur de chaque part. Deux distributions très différentes peuvent avoir les mêmes quartiles.
2. Elles ne se prêtent pas au **calcul algébrique** : l'écart-type intervient dans des formules — la covariance, la corrélation, les intervalles de confiance —, pas l'intervalle interquartile.
3. Elles n'utilisent pas toute l'information disponible.

Le partage des rôles

Quantiles pour décrire et publier; moyenne et écart-type pour calculer et modéliser.

Les deux familles sont complémentaires, et un bon tableau de résultats donne les deux.

Ce qu'il faut retenir

Six points

1. Un **quantile** q_p est la valeur telle qu'une proportion p lui est inférieure : $F(q_p) = p$.
2. Trois quartiles, neuf déciles, quatre-vingt-dix-neuf centiles — et $Q_2 = D_5 = C_{50} = Me$.
3. Sur données groupées, une seule formule d'**interpolation linéaire**, celle de la médiane avec p .
4. La **boîte à moustaches** montre niveau, dispersion, asymétrie et valeurs atypiques d'un seul coup.
5. Règle des $1,5 IQ$: une **convention de repérage**, pas un test.
6. IQ , IQ/Me et D_9/D_1 sont **robustes**; ils décrivent bien mais ne se calculent pas.

La suite

IQ mesure la dispersion à partir des rangs. Il existe une autre famille, fondée sur les **écarts à la moyenne** — et elle donnera l'écart-type, l'indicateur le plus utilisé de toute la statistique.

C'est le chapitre 8.