

Statistique descriptive

Chapitre 3 : Le tableau statistique



Jaouad Madkour

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

Le premier résumé

Les cinq colonnes

Le tableau, en pratique

Les erreurs classiques

Ce qu'il faut retenir

Le premier résumé

Ce qu'on cherche à faire

Les données brutes sont une liste : salarié 1, salarié 2, ... salarié 500. Elle contient toute l'information et n'en délivre aucune.

Le **tableau statistique** regroupe les individus qui partagent la même modalité et ne retient que les **comptages**. On perd l'identité des individus; on gagne une vue d'ensemble.

Le principe

On passe d'une liste de n individus à un tableau de k lignes, où k est le nombre de modalités — et k est toujours **beaucoup plus petit** que n .

Les cinq colonnes

le **tableau statistique** : tout ce qu'on peut dire d'une variable, en cinq colonnes

x_i	n_i	f_i	N_i	F_i
x_1	n_1	f_1	N_1	F_1
x_2	n_2	f_2	N_2	F_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	f_k	$N_k = n$	$F_k = 1$
total	n	1		

x_i les *modalités* ou les classes

n_i l'*effectif* : combien d'individus

$f_i = n_i/n$ la *fréquence* : quelle part

N_i l'*effectif cumulé* : combien au plus

F_i la *fréquence cumulée*

Chaque colonne appelle son propre graphique

— c'est le chapitre 5.

Effectif

L'**effectif** n_i de la modalité x_i est le nombre d'individus qui la possèdent.

$$\sum_{i=1}^k n_i = n$$

Fréquence

La **fréquence** f_i est la part relative de la modalité :

$$f_i = \frac{n_i}{n} \quad \text{avec} \quad 0 \leq f_i \leq 1 \quad \text{et} \quad \sum_{i=1}^k f_i = 1$$

Multipliée par 100, elle donne un **pourcentage**.

Pourquoi les deux colonnes, et pas une seule

L'effectif dit **combien**, la fréquence dit **quelle part**.

L'effectif est indispensable pour juger de la fiabilité : une modalité à 80 % calculée sur 5 individus ne vaut rien. La fréquence est indispensable pour **comparer** deux populations de tailles différentes.

Les deux vérifications à faire systématiquement : la somme des effectifs vaut n , la somme des fréquences vaut 1. Elles détectent la quasi-totalité des erreurs de saisie.

Effectif et fréquence cumulés croissants

$$N_i = \sum_{j \leq i} n_j \qquad F_i = \sum_{j \leq i} f_j$$

N_i répond à : combien d'individus ont une valeur **inférieure ou égale** à x_i ?

Cumuls décroissants

$$N_i^{\downarrow} = n - N_{i-1} \qquad F_i^{\downarrow} = 1 - F_{i-1}$$

Ils répondent à : combien d'individus ont une valeur **supérieure ou égale** à x_i ?

Le cumul est la colonne la plus utile en pratique

Les questions d'un décideur ne portent presque jamais sur une modalité isolée. Elles portent sur des **seuils** :

- combien de salariées gagnent moins de 6 000 dirhams ?
- quelle part de l'effectif dépasse dix ans d'ancienneté ?
- au-dessous de quel salaire se situe la moitié du personnel ?

Ces trois questions se lisent directement dans les colonnes cumulées. La troisième donne la **médiane**, et c'est ainsi qu'on la calculera au chapitre 7.

Rappel du chapitre 2

Cumuler suppose un **ordre**. La colonne N_i n'a donc aucun sens pour une variable nominale : cumuler « Tanger » et « Casablanca » ne veut rien dire.

Elle a un sens pour une variable ordinale, discrète ou continue.

Le tableau, en pratique

Le diplôme de nos 500 salariés

x_i	n_i	f_i	N_i	F_i
Bac	219	0,438	219	0,438
Bac+2	174	0,348	393	0,786
Bac+5	107	0,214	500	1,000
Total	500	1,000		

Ce qu'on lit en trois secondes

Un peu moins de la moitié des salariés s'arrête au baccalauréat. Un salarié sur cinq environ est titulaire d'un Bac+5.

Les cumuls sont ici licites parce que le diplôme est **ordinal** : 78,6 % des salariés ont au plus un Bac+2.

L'ancienneté, valeurs groupées

ancienneté	n_i	f_i	N_i	F_i
0 à 4 ans	82	0,164	82	0,164
5 à 9 ans	180	0,360	262	0,524
10 à 14 ans	139	0,278	401	0,802
15 à 19 ans	66	0,132	467	0,934
20 ans et plus	33	0,066	500	1,000
Total	500	1,000		

La lecture par seuils

Plus de la moitié du personnel — 52,4 % — a moins de dix ans d'ancienneté. Moins d'un salarié sur quatorze a vingt ans ou davantage.

C'est un diagnostic de gestion des ressources humaines obtenu par simple lecture de la colonne F_i : l'entreprise a beaucoup recruté récemment, et sa mémoire interne est concentrée sur peu de personnes.

Les erreurs classiques

Trois pièges à connaître

Piège 1 : les fréquences qui ne somment pas à un

Presque toujours un **arrondi**. $0,438 + 0,348 + 0,214 = 1,000$ ici, mais avec trois modalités à **33,3 %** on obtient **99,9 %**.

Ne jamais forcer le total. On signale l'arrondi en note de bas de tableau : c'est la pratique de tous les instituts statistiques.

Piège 2 : les non-réponses

Si **30** salariés n'ont pas indiqué leur diplôme, faut-il diviser par **500** ou par **470**?

Les deux se pratiquent, mais **il faut le dire**. On distingue :

- la fréquence **brute** : sur l'ensemble, $n = 500$;
- la fréquence **nette** ou « hors non-réponse » : $n = 470$.

Un écart de **6 %** sur le dénominateur est un écart de **6 %** sur tous les résultats.

Piège 3 : le pourcentage sans base

« Les ventes ont progressé de 200 % » : de 1 à 3 unités, ou de 1 million à 3 millions ?

Un pourcentage sans son effectif n'est pas une information. C'est le procédé rhétorique le plus courant dans la presse économique.

La règle professionnelle

Un tableau publié porte toujours, sous le titre :

- le **champ** (quelle population, quelle date);
- l'**effectif total** n ;
- la **source**;
- la mention des non-réponses s'il y en a.

Sans ces quatre éléments, le tableau n'est pas exploitable — même si les chiffres sont exacts.

Ce qu'il faut retenir

Cinq points

1. Le tableau statistique fait passer de n lignes à k lignes : c'est le premier résumé.
2. **Effectif** n_i : combien. **Fréquence** $f_i = n_i/n$: quelle part. Les deux sont nécessaires.
3. **Cumuls** N_i et F_i : ils répondent aux questions de **seuil**, et exigent un ordre.
4. Deux contrôles systématiques : $\sum n_i = n$ et $\sum f_i = 1$.
5. Un tableau se publie avec son **champ**, son **effectif** et sa **source**.

La suite

Pour une variable continue, on ne peut pas énumérer les modalités : il faut les **regrouper en classes**. Et ce regroupement obéit à des règles précises, dont dépend la validité de l'histogramme.

C'est le chapitre 4.