COURS

2de

Tableaux croisés — Fréquences conditionnelles et marginales

Programme de mathématiques de seconde générale et technologique — arrêté du 26 février 2026, BO n° 14 du 2 avril 2026 — MENE2602914A

Un tableau croisé range une population selon deux critères à la fois. Tout le chapitre tient ensuite dans le choix du dénominateur : **une fréquence conditionnelle se calcule dans une sous-population, et « parmi les A, 60 % sont B » ne dit pas la même chose que « parmi les B, 60 % sont A ».** Toutes les données de cette fiche sont fictives.

1 Le tableau croisé d'effectifs

DÉFINITION

Une **variable qualitative** associe à chaque individu d'une population une **modalité** qui n'est pas un nombre. Elle est **nominale** si ses modalités n'ont pas d'ordre (espèce, mode de transport), **ordinale** si elles en ont un (tranche d'âge, degré de satisfaction). Le **tableau croisé d'effectifs** de deux variables qualitatives donne dans chaque **case** le nombre d'individus qui présentent **à la fois** une modalité A de la première et une modalité B de la seconde : c'est l'effectif de l'intersection, $\mathrm{Card}(A \cap B)$. La dernière ligne et la dernière colonne, les **marges**, donnent les totaux ; leur case commune est le **total général**, l'effectif de toute la population.

Port du casque chez 400 cyclistes (données fictives)

âge \ casque	Casque	Sans casque	Total
Moins de 18 ans	60	40	**100**
18 à 40 ans	72	**108**	**180**
Plus de 40 ans	78	42	**120**
Total	**210**	**190**	**400**

une case
108 cyclistes à la fois « 18 à 40 ans » et « sans casque »

les marges
les totaux d'une seule modalité : 100 + 180 + 120 = 210 + 190

le total général
toute la population : 400

Une case : deux conditions réunies. Une marge : une seule condition. Le total général : toute la population.

PROPRIÉTÉ

Les cases d'une ligne ont pour somme le total de la ligne, celles d'une colonne le total de la colonne ; les totaux des lignes, comme ceux des colonnes, ont pour somme le total général. Ces égalités servent à **compléter** un tableau, puis à le **contrôler**.

DRESSER UN TABLEAU CROISÉ À PARTIR D'UN FICHIER

1. **Une ligne par modalité de la première variable, une colonne par modalité de la seconde**, plus une ligne et une colonne « Total ».
2. **Compter chaque case** : les individus qui vérifient la condition de la ligne **et** celle de la colonne — un filtre, au tableur ou en Python avec deux listes et un test en **and**.
3. **Remplir les marges et contrôler** : le total général doit être l'effectif du fichier.

COMPLÉTER UN TABLEAU PAR LES EFFECTIFS

1. **Chercher une ligne ou une colonne, marges comprises, où il ne manque qu'une valeur**, et la déduire par soustraction ou addition.
2. **Recommencer** avec la valeur trouvée, puis **contrôler** : le dernier total, calculé par la ligne et par la colonne, doit concorder.

EXEMPLE

Une enquête interroge 250 clients d'une enseigne. 90 ont acheté en ligne, les autres en magasin ; 200 se disent satisfaits, dont 55 parmi les acheteurs en ligne. Compléter le tableau croisé « mode d'achat × satisfaction ».

1. Ligne « en ligne », puis marge des modes d'achat : il n'y manque qu'une valeur. $90 - 55 = 35$ non-satisfaits en ligne, $250 - 90 = 160$ acheteurs en magasin.
2. Colonne « satisfaits », puis ligne « en magasin » : $200 - 55 = 145$ satisfaits et $160 - 145 = 15$ non-satisfaits en magasin.
3. Contrôle : le total des non-satisfaits, par la marge puis par la colonne. $250 - 200 = 50 = 35 + 15$

En ligne : 55 satisfaits et 35 non satisfaits ; en magasin : 145 et 15 ; au total 200 et 50, pour 250 clients.

PIÈGE

Trouver les satisfaits en magasin par $250 - 55 = 195$. — on a retiré une case du **total général**, qui compte aussi les non-satisfaits : le résultat dépasse même l'effectif des acheteurs en magasin, 160.

Ce que ça coûte : la case, et toutes celles qu'on en déduit ensuite

Le réflexe : soustraire dans la **même ligne** ou la **même colonne** que la case cherchée, et vérifier qu'aucune case ne dépasse sa marge.

2 Fréquences marginales

DÉFINITION

La **fréquence** d'une sous-population dans une population est le quotient de son effectif par l'effectif de la population : un nombre entre 0 et 1, qui s'écrit aussi en fraction ou en pourcentage. Dans un tableau croisé, la **fréquence marginale** d'une modalité est sa fréquence dans **toute la population** : le numérateur se lit dans une **marge**, le dénominateur est le **total général** : $f(A) = \frac{\text{Card}(A)}{\text{effectif total}}$.

CALCULER UNE FRÉQUENCE MARGINALE

1. **Diviser l'effectif lu dans la marge par le total général**, et écrire le résultat sous la forme demandée : décimale, fraction ou pourcentage.
2. **Contrôler** : les fréquences marginales des modalités d'une même variable ont pour somme 1.

EXEMPLE

Pour les 400 cyclistes du tableau ci-dessus, calculer la fréquence marginale des porteurs de casque, puis celles des trois tranches d'âge.

1. Porteurs de casque : 210 sur 400.
   $\frac{210}{400} = 0{,}525$
   soit 52,5 %.
2. Tranches d'âge : $\frac{100}{400} = 0{,}25$, $\frac{180}{400} = 0{,}45$ et $\frac{120}{400} = 0{,}3$.
   $0{,}25 + 0{,}45 + 0{,}3 = 1$

52,5 % des cyclistes portent le casque ; les trois tranches d'âge représentent 25 %, 45 % et 30 % de la population.

PIÈGE

Écrire « la fréquence des porteurs de casque est 0,525 % ». — 0,525 et 52,5 % sont le même nombre ; ajouter % à l'écriture décimale divise la fréquence par cent : 0,525 % de 400 cyclistes, c'est deux cyclistes.

Ce que ça coûte : le résultat, même si le calcul était juste

Le réflexe : multiplier par 100 pour écrire un pourcentage, et contrôler l'ordre de grandeur : 210 sur 400, c'est un peu plus de la moitié.

3 Fréquences conditionnelles

DÉFINITION

Soit A une modalité d'une variable, d'effectif non nul, et B une modalité de l'autre. La **fréquence conditionnelle de B parmi A** — on dit aussi **sachant A** — est la fréquence de B calculée **dans la sous-population A**, qui devient la population de référence. On la note ici $f_A(B)$. Les médias la disent avec « parmi », « chez les », « des … » : « 60 % des moins de 18 ans portent le casque ».

$$f_A(B) = \frac{\text{Card}(A \cap B)}{\text{Card}(A)}$$

parmi les moins de 18 ans

	Casque	Sans casque	Total
Moins de 18 ans	**60**	40	**100**
18 à 40 ans	72	108	180
Plus de 40 ans	78	42	120
Total	210	190	400

60 ÷ 100 = 0,6 = 60 %

60 % des moins de 18 ans portent le casque

parmi les porteurs de casque

	Casque	Sans casque	Total
Moins de 18 ans	**60**	40	100
18 à 40 ans	72	108	180
Plus de 40 ans	78	42	120
Total	**210**	190	400

60 ÷ 210 ≈ 0,286 ≈ 28,6 %

environ 28,6 % des porteurs ont moins de 18 ans

La même case, deux dénominateurs : la population de référence est celle qui suit « parmi ».

CALCULER UNE FRÉQUENCE CONDITIONNELLE

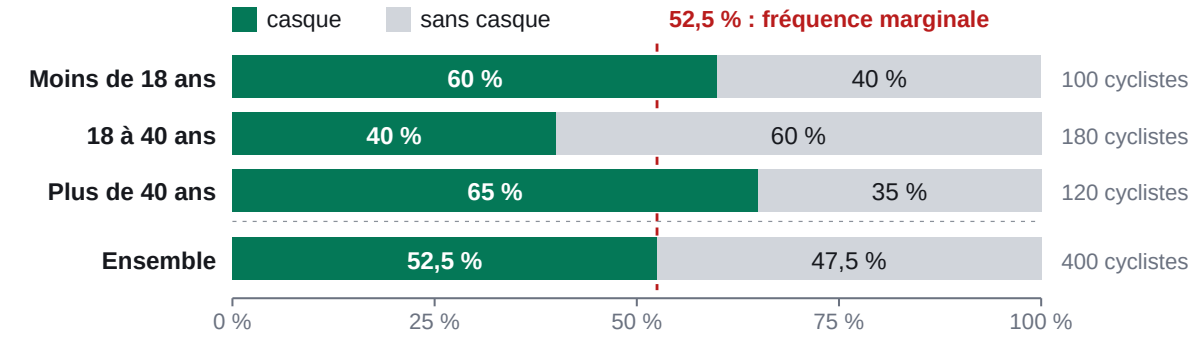
1. **Repérer la population de référence** : le groupe qui suit « parmi », « chez les », « sachant que ». Sa marge fournit le **dénominateur**.
2. **Lire au numérateur la case** qui réunit les deux conditions.
3. **Diviser, arrondir comme demandé, et conclure par une phrase** qui commence par « parmi … » : elle oblige à nommer le bon dénominateur.

EXEMPLE

Toujours pour les 400 cyclistes : calculer la fréquence des porteurs de casque parmi les moins de 18 ans, puis la fréquence des moins de 18 ans parmi les porteurs de casque.

1. Population de référence : les 100 moins de 18 ans, dont 60 portent le casque.
$\frac{60}{100} = 0{,}6$
2. Population de référence : les 210 porteurs de casque, dont 60 ont moins de 18 ans.
$\frac{60}{210} \approx 0{,}286$

Parmi les moins de 18 ans, 60 % portent le casque ; parmi les porteurs de casque, environ 28,6 % ont moins de 18 ans.



Chaque barre est une population de référence ramenée à 100 % : ses fréquences conditionnelles ont pour somme 1. On compare les groupes par ces fréquences, jamais par leurs effectifs ; la fréquence marginale, 52,5 %, tombe entre 40 % et 65 %.

La fréquence marginale est comprise entre les fréquences conditionnelles. Les 210 porteurs de casque sont $60 + 72 + 78$; divisé par 400, chaque terme est le produit d'une fréquence marginale de tranche par une fréquence conditionnelle, par exemple $\frac{72}{400} = \frac{180}{400} \times \frac{72}{180}$. D'où $0{,}525 = 0{,}25 \times 0{,}6 + 0{,}45 \times 0{,}4 + 0{,}3 \times 0{,}65$: une **moyenne pondérée** des fréquences conditionnelles, et non leur moyenne simple. Elle tombe donc entre la plus petite et la plus grande — un contrôle de vraisemblance immédiat.

PIÈGE

De « 60 % des moins de 18 ans portent le casque », conclure que 60 % des porteurs de casque ont moins de 18 ans. — même numérateur, 60, mais pas la même population de référence : 100 jeunes d'un côté, 210 porteurs de casque de l'autre, soit environ 28,6 %.

Ce que ça coûte : la question entière

Le réflexe : commencer la phrase par « parmi … » : ce groupe fournit le dénominateur.

PIÈGE

Affirmer que les 18 à 40 ans portent plus souvent le casque que les moins de 18 ans, parce que 72 > 60. — les groupes n'ont pas le même effectif : 72 sur 180 font 40 %, 60 sur 100 font 60 %. C'est l'inverse.

Ce que ça coûte : la conclusion, qui est le cœur de la question

Le réflexe : comparer des groupes de tailles différentes par leurs fréquences conditionnelles.

PIÈGE

Obtenir la fréquence globale des porteurs de casque par $\frac{60+40+65}{3} = 55\,\%$. — elle donne le même poids à 100 jeunes et à 180 adultes ; la vraie valeur est $\frac{210}{400}$, soit 52,5 %.

Ce que ça coûte : le résultat, et toute comparaison qui s'appuie dessus

Le réflexe : revenir aux effectifs : additionner les numérateurs, additionner les dénominateurs, puis diviser.

4 Compléter un tableau avec des fréquences conditionnelles

RÈGLE

Une fréquence conditionnelle relie trois nombres. Elle donne **une partie connaissant le tout** — la case, connaissant l'effectif de la sous-population — et, en divisant, **le tout connaissant une partie**.

$$\mathrm{Card}(A \cap B) = f_A(B) \times \mathrm{Card}(A) \qquad \mathrm{Card}(A) = \frac{\mathrm{Card}(A \cap B)}{f_A(B)}$$

COMPLÉTER UN TABLEAU À PARTIR DE POURCENTAGES

1. **Traduire chaque pourcentage** : « 70 % des cadres » s'applique à l'effectif des cadres, « 40 % des salariés » au total général.
2. **Calculer les marges connues, puis les cases** par $f_A(B) \times \text{Card}(A)$; finir par soustractions et additions, et contrôler les marges.

EXEMPLE

Une entreprise compte 800 salariés, dont 40 % de cadres. Parmi les cadres, 70 % pratiquent le télétravail ; parmi les non-cadres, 25 %. Compléter le tableau croisé, puis calculer la fréquence des cadres parmi les salariés en télétravail.

1. Marges : $0{,}4 \times 800 = 320$ cadres et $800 - 320 = 480$ non-cadres. Puis chaque case « télétravail », avec l'effectif de **sa** sous-population :
   $0{,}7 \times 320 = 224$
   et $0{,}25 \times 480 = 120$.
2. Le reste par additions et soustractions : $224 + 120 = 344$ en télétravail, $320 - 224 = 96$ et $480 - 120 = 360$ sans.
   $344 + 96 + 360 = 800$
   contrôle par le total général.
3. Population de référence : les 344 salariés en télétravail, dont 224 cadres.
   $\frac{224}{344} \approx 0{,}651$

Parmi les salariés en télétravail, environ 65,1 % sont des cadres — à ne pas confondre avec les 70 % de cadres qui télétravaillent.

800 salariés (données fictives) : en noir la donnée, en vert les calculs

	Télétravail	Sans télétravail	Total
Cadres	224 (2)	96 (5)	320 (1)
Non-cadres	120 (3)	360 (5)	480 (1)
Total	344 (4)	456 (5)	800

1. 40 % de 800 = 320, reste 480
2. 70 % des 320 cadres = 224
3. 25 % des 480 non-cadres = 120
4. 224 + 120 = 344
5. les autres : soustractions

Les marges d'abord, puis les cases par « pourcentage × effectif de la sous-population », enfin les soustractions.

PIÈGE

Calculer les cadres en télétravail par $0{,}7 \times 800 = 560$. — 70 % s'applique aux cadres, pas à tous les salariés : la « partie » obtenue, 560, dépasse même le nombre de cadres, 320.

Ce que ça coûte : la case, et tout le tableau qui en découle

Le réflexe : avant de multiplier, nommer la population qui suit « parmi » et prendre son effectif.

Vers les probabilités conditionnelles. Tirons au sort un salarié, avec équiprobabilité. La probabilité qu'il soit cadre, $P(C) = \frac{\text{Card}(C)}{\text{Card}(\Omega)} = \frac{320}{800}$, est la fréquence marginale ; la probabilité qu'il télétravaille **sachant** qu'il est cadre, $P_C(T) = \frac{\text{Card}(C \cap T)}{\text{Card}(C)} = \frac{224}{320}$, est la fréquence conditionnelle. La fiche Probabilités prolonge ce lien avec les arbres.

À VÉRIFIER

- Chaque ligne et chaque colonne ont-elles pour somme leur total, et les marges le total général ?
- Pour chaque fréquence, ai-je nommé la **population de référence** avant de diviser ?
- Ai-je distingué « parmi les A, … sont B » de « parmi les B, … sont A » ?
- Ai-je comparé des groupes de tailles différentes par des **fréquences**, et non par des effectifs ?

À RETENIR

- **Case** : les deux conditions à la fois ; **marge** : une seule ; **total général** : toute la population.
- Fréquence marginale = $\frac{\text{marge}}{\text{total général}}$.
- Fréquence de B parmi A : $f_A(B) = \frac{\text{Card}(A \cap B)}{\text{Card}(A)}$ — le dénominateur est le groupe qui suit « parmi ».
- Une case vaut $f_A(B) \times \text{Card}(A)$: le pourcentage s'applique à sa sous-population.
- On compare des groupes par leurs fréquences conditionnelles ; la fréquence globale tombe entre elles.