

Fiche de cours

Tableaux croisés et fréquences conditionnelles

L'essentiel du chapitre en deux pages

définitions · propriétés · méthodes · à savoir par cœur

1 LE TABLEAU CROISÉ D'EFFECTIFS

Définition. Un **tableau croisé** répartit les N individus d'une population selon **deux** caractères. Chaque case donne l'effectif des individus qui possèdent à la fois la modalité de sa ligne et celle de sa colonne. La dernière colonne donne les **totaux-lignes**, la dernière ligne les **totaux-colonnes**, et la case du coin l'**effectif total** N .

Propriété. Chaque total-ligne est la somme de sa ligne, chaque total-colonne la somme de sa colonne, et la somme des totaux-lignes est égale à la somme des totaux-colonnes, égale à N . C'est ce qui permet de **compléter** un tableau dont des cases manquent, et de vérifier ses calculs.

Effectifs	Moins de 30 ans	30 ans ou plus	Total
Vélo	150	90	240
Voiture	120	240	360
Total	270	330	600

Tableau 1 — Mode de déplacement domicile-travail de 600 personnes, selon l'âge. Ce tableau sert d'exemple dans toute la fiche.

2 FRÉQUENCES ET FRÉQUENCES MARGINALES

Définition. La **fréquence** d'un groupe est le quotient de son effectif par l'effectif total N ; on l'exprime souvent en pourcentage. Les fréquences des cases forment le **tableau des fréquences** ; leur somme vaut 1, soit 100 %.

Définition. Les fréquences obtenues à partir des **totaux** d'une ligne ou d'une colonne sont les **fréquences marginales** : elles décrivent un seul des deux caractères.

Remarque. Dans le tableau 1 : la fréquence marginale du vélo est $\frac{240}{600} = 0,40$, soit 40 % ; celle des moins de 30 ans est $\frac{270}{600} = 0,45$, soit 45 %. La fréquence de la case « vélo et moins de 30 ans » est $\frac{150}{600} = 0,25$, soit 25 %.

Méthode — des effectifs aux fréquences, et réciproquement.

Pour obtenir le tableau des fréquences, diviser *toutes* les cases par N ; pour revenir aux effectifs, multiplier toutes les fréquences par N . Une fréquence marginale est aussi la somme des fréquences de sa ligne (ou de sa colonne) : c'est un moyen de contrôle.

Fréquences	Moins de 30 ans	30 ans ou plus	Total
Vélo	25 %	15 %	40 %
Voiture	20 %	40 %	60 %
Total	45 %	55 %	100 %

Tableau 2 — Le tableau des fréquences associé au tableau 1. Les fréquences marginales sont dans la dernière ligne et la dernière colonne.

3 FRÉQUENCES CONDITIONNELLES

Définition. La **fréquence conditionnelle** de A parmi les B (on dit aussi « sachant que » l'individu est un B) est

$$f_B(A) = \frac{n(A \text{ et } B)}{n(B)} : \text{ le dénominateur est le total du groupe } B, \text{ et non } N.$$

Méthode — repérer le bon dénominateur. Le mot qui suit « parmi », « sachant que », « lorsque », « chez les » désigne le groupe de *référence* : c'est lui qui fournit le dénominateur. On travaille alors dans une seule ligne (ou une seule colonne) du tableau.

Propriété. Pour un groupe de référence fixé, les fréquences conditionnelles ont pour somme 1 : $f_B(A) + f_B(\text{non } A) = 1$. En revanche $f_B(A)$ et $f_A(B)$ sont en général **différentes**.

Remarque. Dans le tableau 1 : parmi les moins de 30 ans, la fréquence du vélo est $\frac{150}{270} \approx 0,556$, soit environ 55,6 % ; parmi les cyclistes, la fréquence des moins de 30 ans est $\frac{150}{240} = 0,625$, soit 62,5 %. Les deux phrases ne disent pas la même chose.

4 « ET », « OU », « NON » : COMPTER SUR UN TABLEAU

Propriété. Pour deux groupes A et B d'une même population de N individus :

■ FORMULAIRE

- « A et B » : c'est une case du tableau croisé.
- $n(\text{non } A) = N - n(A)$ (le groupe **complémentaire**).
- $n(A \text{ ou } B) = n(A) + n(B) - n(A \text{ et } B)$: on retranche la case comptée deux fois.
- Les fréquences correspondantes s'obtiennent en divisant par N .

```
1 n = 0
2 for (age, mode) in donnees:
3     if age < 30 and mode == "velo":
4         n = n + 1
5 print(n, n/len(donnees))
```

Programme 1 — Effectif et fréquence du groupe « moins de 30 ans et vélo » ; donnees contient un couple (âge, mode) par personne. Remplacer and par or donne le groupe « ou ».

Méthode — filtrer des données. Dans un tableur, un comptage conditionnel s'obtient avec NB.SI (une condition) ou NB.SI.ENS (plusieurs conditions combinées par « et »). En Python, on parcourt les données et on compte les individus qui vérifient la condition, écrite avec and, or et not.

5 SAVOIR-FAIRE DU CHAPITRE

➤ SAVOIR-FAIRE

- Lire un tableau croisé et le compléter à partir d'effectifs ou de fréquences partiels.
- Calculer une fréquence, une fréquence marginale, et construire le tableau des fréquences.
- Calculer une fréquence conditionnelle et l'énoncer par une phrase sans ambiguïté ; distinguer $f_A(B)$ et $f_B(A)$.
- Traduire « et », « ou », « non » sur un tableau croisé et appliquer $n(A \text{ ou } B) = n(A) + n(B) - n(A \text{ et } B)$.
- Passer d'un tableau de fréquences à un tableau d'effectifs ; interpréter un comptage obtenu par filtre (tableur ou Python).

6 EXEMPLES TYPES

★ EXEMPLE — fréquence marginale et fréquence conditionnelle

Avec le tableau 1 : la fréquence des personnes en voiture est $\frac{360}{600} = 60\%$ (fréquence marginale). Parmi les 330 personnes de 30 ans ou plus, 240 sont en voiture, soit $\frac{240}{330} \approx 72,7\%$: la voiture est nettement plus fréquente dans cette tranche d'âge.

★ EXEMPLE — « ou » et « non »

Toujours avec le tableau 1 : $n(\text{vélo ou moins de 30 ans}) = 240 + 270 - 150 = 360$, soit une fréquence de $\frac{360}{600} = 60\%$. Et $n(\text{non vélo}) = 600 - 240 = 360$: c'est bien le total de la ligne « voiture ».

★ EXEMPLE — compléter un tableau à partir de fréquences

Une enquête porte sur 400 personnes, dont 55 % de femmes ; parmi les femmes, 20 % pratiquent la natation. On en déduit : $0,55 \times 400 = 220$ femmes, puis $0,20 \times 220 = 44$ femmes nageuses. La fréquence des femmes nageuses dans l'ensemble est $\frac{44}{400} = 11\%$, et non 20 %.

✗ PIÈGES

- Diviser par N au lieu du total du groupe de référence : une fréquence conditionnelle a pour dénominateur un *total-ligne* ou un *total-colonne*.
- Confondre « 62,5 % des cyclistes ont moins de 30 ans » (vrai dans le tableau 1) et « 62,5 % des moins de 30 ans sont cyclistes » (faux : environ 55,6 %) : les deux fréquences n'ont pas le même sens et, en général, pas la même valeur.
- Additionner les effectifs de A et de B pour obtenir « A ou B » sans retrancher la case « A et B ».
- Additionner (ou moyenner) des fréquences calculées sur des groupes de référence différents : elles n'ont pas le même dénominateur, leur somme n'a pas de sens.
- Arrondir avant la fin : conserver les effectifs exacts et n'arrondir qu'au dernier calcul.