

# Chapitre 3 : Tableaux statistiques croisés et probabilités conditionnelles

## 1) Tableaux croisés, marges, effectifs marginaux

### Définitions :

Un tableau croisé est un tableau à double entrée qui permet de donner la répartition d'une population en fonction de **deux caractères**.

**La ligne « Total » et la colonne « Total » sont appelées les marges** du tableau croisé.

### Exemple :

Suite à une enquête auprès de 1500 clients, le tableau suivant donne la répartition des effectifs des clients d'un fournisseur internet, selon leur niveau de satisfaction et la formule de l'abonnement souscrit :

	Pas satisfait	Satisfait	Total
Abonnement Basique	158	622	780
Abonnement Ciné+	41	257	298
Abonnement Foot+	69	353	422
Total	268	1232	1500

Remarque : chacune des marges donne la répartition de la population selon un caractère (niveau de satisfaction pour la ligne « total », formule de l'abonnement souscrit pour la colonne « total »), les effectifs indiqués dans les marges sont appelés effectifs marginaux.

## 2) Fréquences, Fréquences marginales

### Définitions :

Un tableau croisé est un tableau à double entrée qui permet de donner la répartition d'une population en fonction de **deux caractères**.

**La ligne « Total » et la colonne « Total » sont appelées les marges** du tableau croisé.

Reprenons l'exemple précédent et calculons les fréquences par rapport à l'effectif total :

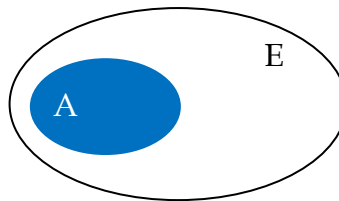
Rappel : fréquence (ou proportion) d'une sous population A dans une population E :

$$f(A) = \frac{n_A}{n_E}$$

avec  $n_A$  effectif de la sous-population A et  $n_E$  effectif de la population E.

On peut aussi écrire :

$$f(A) = \frac{\text{card}\{A\}}{\text{Card}\{E\}}$$



	Pas satisfait	Satisfait	Total
Abonnement Basique (B)	$\frac{158}{1500} \approx$	$\frac{622}{1500} \approx$	$\frac{780}{1500} \approx$
Abonnement Ciné+ (C)	$\frac{41}{1500} \approx$	$\frac{257}{1500} \approx$	$\frac{298}{1500} \approx$
Abonnement Foot+ (F)	$\frac{69}{1500} \approx$	$\frac{353}{1500} \approx$	$\frac{422}{1500} \approx$
Total	$\frac{268}{1500} \approx$	$\frac{1232}{1500} \approx$	1 (100%)

### 3) Fréquences conditionnelles

Approche :

Chaque caractère permet de partager la population globale en sous-populations, par exemple :

- la sous-population des clients non satisfaits ( $\bar{S}$ ) et celle des clients satisfaits ( $S$ ),
- les sous-populations respectives des clients ayant choisi l'abonnement Basic ( $B$ ), Ciné+ ( $C$ ), et Foot ( $F$ ).

⇒ Complétons le tableau suivant, dit **tableau des fréquences conditionnelles en ligne**, afin de comparer la satisfaction des clients en fonction du type d'abonnement :

	Pas satisfait	Satisfait	Total
Abonnement Basique (B)	$\frac{158}{780} \approx$	$\frac{622}{780} \approx$	$\frac{780}{780} = 1 (= 100\%)$
Abonnement Ciné+ (C)	$\frac{41}{298} \approx$	$\frac{257}{298} \approx$	$\frac{298}{298} = 1 (= 100\%)$
Abonnement Foot+ (F)	$\frac{69}{422} \approx$	$\frac{353}{422} \approx$	$\frac{422}{422} = 1 (= 100\%)$

### Exemple de calcul de fréquences conditionnelles en colonne :

Une association de service à domicile a fait une étude statistique sur les malades de la grippe parmi ses salariés sur la période du 1<sup>er</sup> septembre 2019 au 1<sup>er</sup> mai 2020 en tenant également en compte le critère de vaccination. La répartition des effectifs est donnée dans le tableau suivant :

	Vaccinés	Non vaccinés	Total
Malade	25	27	52
Non malade	552	206	758
Total	577	233	810

⇒ Complétons le tableau suivant, dit **tableau des fréquences conditionnelles en colonne**, afin de comparer la fréquence de malades parmi les salariés vaccinés et non vaccinés :

	Vaccinés	Non vaccinés
Malade		
Non malade		
Total	<b>1(=100%)</b>	<b>1 (100%)</b>



## **4) Probabilités conditionnelles**

En statistiques nous calculons des fréquences à partir de données observées. Les probabilités prédisent des résultats d'expériences aléatoires (régies par le hasard). **Les exemples précédents peuvent être transposés en calcul de probabilités si l'on introduit « un tirage au sort ».**

Exemples : on tire au sort un salarié,

- Quelle est la probabilité de l'évènement  $M$  : « le salarié est malade » ?  
 $P(M) = \dots\dots\dots$
- Quelle est la probabilité de l'évènement  $M \cap V$  : « le salarié est malade et vacciné » ?  
 $P(M \cap V) = \dots\dots\dots$
- Quelle est la probabilité de l'évènement : « le salarié est Malade SACHANT QU'il est Vacciné » que l'on note  $P_V(M)$  ?  
 $P_V(M) = \dots\dots\dots$

D'une manière générale :

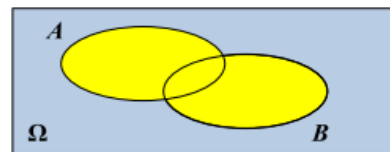
Soit  $A$  et  $B$  deux évènements tels que  $\text{card}\{A\} \neq 0$  et donc  $P(A) \neq 0$   
 La probabilité que l'évènement  $B$  se produise sachant que l'évènement  $A$  est réalisé est donnée par :

$$P_A(B) = \frac{\text{card}\{A \cap B\}}{\text{Card}\{A\}} \text{ et } P_A(B) = \frac{P(A \cap B)}{P(A)}$$

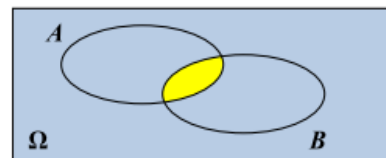
## 5) Rappel : probabilités d'une union d'évènements

### Propriété :

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



$A \cup B$



$A \cap B$

### Exemple d'application :

Afin de suivre l'évolution d'une pandémie au sein de la population française et pouvoir ainsi étudier l'immunité collective, on a établi le tableau suivant en ramenant les données sur un total de 10000 individus :

	Vaccinés	Non vaccinés	Total
Malade*	1284	9369	10653
Non malade	64594	24574	89168
Décédés	23	156	179
Total	65901	34099	100 000

*\*Malade au cours de 6 derniers mois*

On désignera par  $G$  et  $V$  les évènements suivants :

$G$  : « l'individu est guéri »

$V$  : « l'individu est vacciné »

On choisit un individu au hasard parmi la population, quelle est la probabilité qu'il soit guéris ou vacciné ?

.....

.....

.....

.....