❉ Chapitre 1 ❉

Croisement de deux variables

I. Rappel de seconde

1. Proportion sur un ensemble

Définition 1:

Soit A une partie d'une population E. La proportion des éléments de A **par rapport à** E est

$$p = \frac{\text{Nombre d'éléments de A}}{\text{Nombre d'éléments de E}} = \frac{n_A}{n_E}$$

Exemple 1:

Dans une population de référence constitué de 400 personnes, on suppose que 56 personnes ont une particularité A. La proportion de personnes ayant la particularité A est $\frac{56}{400}$. Pour exprimer cette proportion en un pourcentage, on peut procéder comme suit :

$$\frac{56}{400} = \frac{56 \times 100}{400 \times 100} = \frac{56}{4} \times \frac{1}{100} = 14 \times \frac{1}{100} = 14\%$$

Remarque :

On voit qu'une proportion peut s'exprimer sous trois formes différentes. Si on reprend l'exemple précédent, on peut exprimer la proportion de personnes ayant la particularité A sous forme :

- décimale : $\frac{56}{400} = 0.14$
- d'une fraction irréductible : $\frac{56}{400} = \frac{7}{50}$
- d'un pourcentage : $\frac{56}{400} = 14\%$

2. Relation d'inclusion

Définition 2:

Un ensemble A est inclus dans un ensemble B, noté $A \subset B$, lorsque tous les éléments de A appartiennent à B.

Notons que si A est une sous-population d'une population E alors naturellement, $A \subset E$.

Exemple 2:

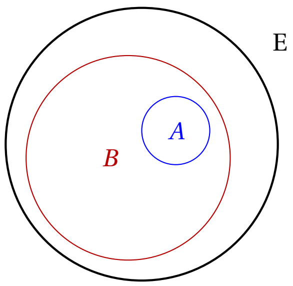
Soit E l'ensemble des élèves du lycée et A les élèves de cette classe de 1STMG. On voit que tous les élèves de 1STMG sont aussi des élèves du lycée. On peut donc en déduire que $A \subset E$.

Propriété 1 :

Dans une population E, considérons deux sous-populations A et B. Supposons que A est incluse dans B, alors $A \subset B \subset E$:

De plus, la proportion p_A de A dans E est égale au produit de la proportion p' de A dans B et de la proportion p_B de B dans E :

$$\boxed{p_A = p' \times p_B}$$



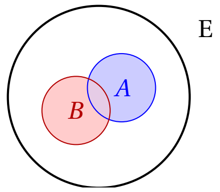
Exemple 3:

D'après les chiffres donné par l'administration, la proportion p_1 d'élèves du lycée en 1STL est de 3.7%, eux-même répartis en deux classes. Donc, la proportion p_2 d'élèves de 1STL dans cette classe est de : $p_2 = \frac{1}{2}$.
La proportion p d'élèves du lycée dans cette classe est donc de :

$$\begin{aligned} p &= p_1 \times p_2 \\ &= \frac{3.7}{100} \times \frac{1}{2} \\ &= \frac{3.7}{200} = 0.0185 = 1.85\% \end{aligned}$$

Il y a donc 1.85% des élèves du lycée qui sont dans cette classe.

II. Croisement de deux variables



Soit A et B deux parties (ou sous populations) d'une population E.
On peut représenter cela par un diagramme de Venn comme ci-contre :

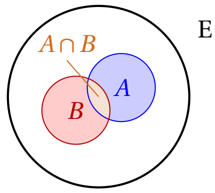
Exemple 4:

On considère comme population E les élèves de la classe de 1STMG et on désigne par A ceux qui sont des garçons, par B ceux qui ont une calculatrice graphique et par C ceux qui sont des filles.

1. Intersection

Définition 3:

L'intersection des parties A et B, noté $A \cap B$ (lire « A inter B »), est constitué des éléments qui sont dans A **et** dans B.



Exemple 5:

En reprenant l'exemple précédent, l'ensemble $A \cap B$ est composé des personnes qui sont dans A **et** dans B *(en même temps)*. On peut donc dire que ce sont les élèves de sexe masculin de 1STL qui possèdent une calculatrice graphique.

Définition 4:

Deux sous-populations A et B d'une même population E sont **disjointes** lorsqu'elles n'ont pas d'élément commun et on note : $A \cap B = \emptyset$

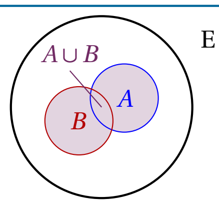
Exemple 6:

En reprenant l'exemple précédent, on voit que $A \cap B \neq \emptyset$ car il y a des élèves de sexe masculin de 1STL qui possèdent une calculatrice graphique. Par contre, on voit que $A \cap C = \emptyset$ car aucun élèves n'est à la fois une fille ET un garçon. Ces deux ensembles sont donc **disjoints**.

2. Réunion

Définition 5:

La réunion des parties A et B, noté $A \cup B$ (lire « A union B », est constitué des éléments qui sont dans A **ou** dans B.



Exemple 7:
En reprenant l'exemple précédent, l'ensemble $A \cup B$ est composé des personnes qui sont dans A **ou** dans B *(ou dans les deux)*. On peut donc dire que ce sont les élèves de 1STL de sexe masculin ou les élèves qui possèdent une calculatrice graphique *(ou dans les deux)*.

Propriété 2 :
Soit A et B deux sous populations d'une même population E. On note p_A la proportion d'individus de A dans E et de même, p_B, $p_{A \cap B}$, $p_{A \cup B}$. On a la relation suivante :

$$p_{A \cup B} = p_A + p_B - p_{A \cap B}$$

Exemple 8:
En reprenant l'exemple précédent, on sait que A et C sont disjoints *(car $A \cap C = \emptyset$)* donc que $p_{A \cap C} = 0$. D'après la propriété précédente, on obtient dans ce cas :

$$\begin{array}{rcl} p_{A \cup C} & = & p_A + p_C - p_{A \cap C} \\ & = & p_A + p_C - 0 \\ & = & p_A + p_C \end{array}$$

III. Tableaux croisés

1. Un exemple pour comprendre

Exemple 9:
On s'intéresse à deux caractères d'une population, par exemple le nombre d'admis au diplôme du baccalauréat en 2018 en fonction de la filière du candidat. On peut alors les représenter à l'aide d'un tableau croisé.

	Admis	Recalé	Total
Filière STL	8 449	789	9 238
Filière STMG	63 690	10 396	74 086
Filière S	187 629	17 109	204 738
Total	259 768	28 294	288 062

L'exemple précédent s'intéresse à trois sous-populations ayant deux modalités. On pourrait tout à fait imaginé que les caractères A et B étudiés possèdent n-modalités. Le tableaux serait donc sous la forme :

	A_1	...	A_n	Total
B_1	Effectif de $A_1 \cap B_1$	...	Effectif de $A_n \cap B_1$	Effectif de B_1
...	...	...	...	...
B_n	Effectif de $A_1 \cap B_n$	...	Effectif de $A_n \cap B_n$	Effectif de B_n
Total	Effectif de A_1	...	Effectif de A_n	Effectif de total

Les deux colonne et ligne « Total » sont appelées les marges du tableau.

2. Fréquences marginales

On divise chaque case du tableau par l'effectif total : on obtient ainsi dans les marges les fréquences dites marginales.

Exemple 10:

On reprend le tableau d'effectifs de l'exemple précédent. L'effectif total vaut 288 062. Donc on divise tous les nombres du tableau par cette valeur pour obtenir les fréquences marginales.

	(A) Admis	(R) Recalé	Total
(L) Filière STL	$\frac{8\,449}{288\,062} \simeq 0{,}0293$	$\frac{789}{288\,062} \simeq 0{,}0027$	$\frac{9\,238}{288\,062} \simeq 0{,}0321$
(G) Filière STMG	$\frac{63\,690}{288\,062} \simeq 0{,}2211$	$\frac{10\,396}{288\,062} \simeq 0{,}0361$	$\frac{74\,086}{288\,062} \simeq 0{,}2572$
(S) Filière S	$\frac{187\,629}{288\,062} \simeq 0{,}6513$	$\frac{17\,109}{288\,062} \simeq 0{,}0594$	$\frac{204\,738}{288\,062} \simeq 0{,}7107$
Total	$\frac{259\,768}{288\,062} \simeq 0{,}9018$	$\frac{28\,294}{288\,062} \simeq 0{,}0982$	$\frac{288\,062}{288\,062} = 1$

On a $f(L \cap A) = 2{,}93\%$. Donc les bacheliers STL ont donc une fréquence de 0.0293 soit $2{,}93\%$ de l'ensemble étudié.

3. Fréquences conditionnelles

Par lignes

On divise l'effectif de chaque case par l'effectif totale de la ligne correspondante. On parle de fréquence conditionnelle car l'ensemble de référence pour le calcul des fréquences n'est plus celui de la population totale : ce sera successivement l'ensemble des STL, des STMG, des S et enfin de la population totale.

Exemple 11:

	(A) Admis	(R) Recalé	Total
(L) Filière STL	$\frac{8\,449}{9\,238} \simeq 0.915$	$\frac{789}{9\,238} \simeq 0{,}085$	$\frac{9\,238}{9\,238} = 1$
(G) Filière STMG	$\frac{63\,690}{74\,086} \simeq 0{,}86$	$\frac{10\,396}{74\,086} \simeq 0{,}14$	$\frac{74\,086}{74\,086} = 1$
(S) Filière S	$\frac{187\,629}{204\,738} \simeq 0{,}916$	$\frac{17\,109}{204\,738} \simeq 0{,}084$	$\frac{204\,738}{204\,738} = 1$

$f_L(A) = \dfrac{f(L \cap A)}{f(L)} = 0.915.$ Le taux de réussite au baccalauréat est de $91{,}5\%$ en filière STL.

$f_G(R) = \dfrac{f(G \cap R)}{f(G)} = 0.14.$ Le taux d'échec au baccalauréat est de 14% en filière STMG.

Par colonnes

On procède comme pour les lignes sauf que l'on divise par l'effectif total de la colonne.

Exemple 12:

$f_A(L) = \dfrac{f(L \cap A)}{f(A)} = 0.0325.$ Parmi les bacheliers, il y a $3{,}25\%$ de STL.

$f_R(S) = \dfrac{f(S \cap R)}{f(R)} = 0.6047.$ Parmi les recalés, il y a $60{,}47\%$ de S.