

Tableau de contingence

I Vocabulaire

Population : ensemble sur lequel porte l'étude.

Individu : un élément de cet ensemble.

Échantillon : une partie de la population.

Recensement : enquête qui porte sur toute la population.

Sondage : enquête qui porte sur une partie de la population.

Caractère : critère qui permet de partager la population en réalisant une partition* de l'ensemble.

* Soit E un ensemble non vide. Une **partition** de E est un ensemble de parties de E qui sont non vides, disjointes deux à deux et dont la réunion est égale à E.

Le caractère étudié dans la population peut-être :

qualitatif nominal : la couleur des yeux,
 ordinal : la puissance fiscale d'un véhicule.

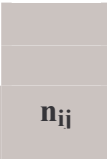

quantitatif discret : le nombre d'enfants par famille
 continue : la taille, le poids.

II Tableaux de contingence

Voici quelques éléments d'information sur les notations et le vocabulaire concernant les tableaux de contingence. Il n'est pas question d'introduire certaines de ces notions ou notations dans un cours de baccalauréat professionnel.

A) Tableau d'effectifs pour deux caractères

On note r le nombre de lignes (nombre de modalités du caractère A) et s le nombre de colonnes (nombre de modalités du caractère B).

B	b₁	...	b_j	...	b_s	Marge de A	
A							
a₁					n_{1.}		
...							
a_i					n_{i.}		
...							
a_r					n_{r.}		
Marge de B	n_{.1}				n_{.j}	n_{.s}	n

n_{ij} : i pour le numéro de la ligne, puis j pour celui de la colonne.

n_{ij} est le nombre d'individus qui possèdent la modalité a_i du caractère A et la modalité b_j du caractère B. Les effectifs lus à l'intérieur du tableau de contingence sont appelés les effectifs partiels.

$n_{i.}$ est le nombre d'individus qui possèdent la modalité a_i du caractère A.

$n_{.j}$ est le nombre d'individus qui possèdent la modalité b_j du caractère B.

Les effectifs lus dans les deux marges du tableau sont appelés effectifs marginaux.

B) Tableau des fréquences pour deux caractères

- **Fréquences partielles** : les fréquences partielles sont obtenues en faisant le rapport des effectifs partiels à l'effectif total.
La proportion d'individus possédant à la fois la modalité a_i du caractère A et la modalité b_j du caractère B est notée f_{ij} , on a donc : $f_{ij} = n_{ij}/n$
- **Fréquences marginales** : les fréquences marginales sont obtenues en faisant le rapport des effectifs marginaux à l'effectif total.
La proportion d'individus possédant la modalité a_i du caractère A est notée $f_{i.}$ et on a :
 $f_{i.} = n_{i.}/n$
- **Fréquences conditionnelles** : les fréquences conditionnelles sont obtenues en faisant le rapport de l'effectif partiel à l'effectif marginal.
La proportion d'individus possédant la modalité a_i du caractère A sous condition qu'ils possèdent la modalité b_j du caractère B est notée $f_{i/j}$ et on a : $f_{i/j} = n_{ij}/n_{.j}$.

Tableau des fréquences partielles

		B					Marge de A
		b_1	...	b_j	...	b_s	
A							
	a_1						$f_{1.}$
	...						
	a_i			f_{ij}			$f_{i.}$
	...						
	a_r						$f_{r.}$
	Marge de B	$f_{.1}$		$f_{.j}$		$f_{.s}$	1

Ces notations ne sont pas à introduire dans un cours de baccalauréat professionnel. Cependant il est possible de demander un calcul de fréquence partielle (ou de pourcentage) ou de fréquence conditionnelle (ou de pourcentage) à partir d'une phrase clairement exprimée (voir les exemples donnés plus loin).

Vocabulaire:

Tri à plat : on entend par cette expression un tri de la population selon un seul caractère.

Tri croisé : on trie la population suivant deux caractères.

Un tableau de contingence est un tableau des effectifs de la population triée selon deux caractères.

III Exemple et contre-exemple

Exemple :

Le tableau suivant est tiré de la revue « TEPC 1996 » publiée par l'INSEE Poitou-Charentes. On considère les immatriculations, en 1995, des véhicules de la région Poitou-Charentes.

	B	Charente	Charente-Maritime	Deux-Sèvres	Vienne
A					
Voitures neuves		9724	15939	8703	10859
Voitures d'occasion		28460	43794	29290	30320
Cars et autobus neufs		23	23	32	17
Cars et autobus d'occasion		26	43	51	43
Camionnettes, camions neufs		1824	2712	1805	8326
Camionnettes, camions d'occasion		5107	8499	5004	5239

Voici une liste, non exhaustive, de questions que l'on peut poser à partir de ce tableau :

- Que signifie le nombre 29290 ?
- Dans le département de la Charente, quel est le pourcentage des immatriculations de camionnettes et camions neufs par rapport au parc des véhicules de la Charente ?
- Dans quel département le pourcentage d'immatriculations de camionnettes ou camions neufs est-il le plus faible ?
- Donner la répartition en pourcentages des immatriculations de véhicules dans le département de la Vienne.

Exemple d'un tableau qui n'est pas un tableau de contingence.

Rappel : dans un tableau de contingence chaque caractère doit permettre d'effectuer une partition de l'ensemble de la population.

L'individu considéré dans le tableau suivant est la commune.

	B	Charente	Charente-Maritime	Deux-Sèvres	Vienne
A					
École de musique		15	57	53	55
Musée		14	39	18	20
Bibliothèque		73	111	88	101

Les différentes modalités de la variable A, équipement culturel, ne permettent pas de faire une partition des communes (une commune peut à la fois avoir son école de musique, sa bibliothèque ainsi que son musée), même si les modalités de la variable B permettent une partition. Ce tableau n'est pas un tableau de contingence !

Par contre le tableau du premier exemple est un tableau de contingence. Un véhicule est immatriculé dans un seul département et n'appartient qu'à une seule catégorie de véhicule.

IV Exercices

Exercice 1

A Sexe	Masculin	Féminin	Marge de A
B C.S.P.			
Agriculteur	1200	600	
Cadre supérieur	1400	500	
Cadre moyen	1800	1300	
Employés	1500	3200	
Ouvriers	4500	2300	
Marge de B			#

- 1° Quels sont les caractères étudiés pour chaque individu ?
- 2° Que représente l'effectif 1800 du tableau ?
- 3° Complétez les marges du tableau (lignes et colonnes vides).
- 4° Que représente la case marquée # ?
- 5° Citez d'autres caractères (au moins 4) qui auraient pu permettre de partitionner cette population. Dites pour chacun d'eux si le caractère est qualitatif ou quantitatif.

Corrigé :

- 1° Les caractères étudiés sont le sexe et la catégorie socio-professionnelle.
- 2° Le nombre 1800 est le nombre de cadres moyens masculins.
- 3° Les nombres inscrits dans la marge de droite représentent respectivement le nombre d'agriculteurs, de cadres supérieurs, de cadres moyens, d'employés et d'ouvriers.
Les nombres inscrits dans la marge du bas du tableau représentent respectivement le nombre d'hommes et de femmes.

A Sexe	Masculin	Féminin	Marge de A
B Métier			
Agriculteur	1200	600	1800
Cadre supérieur	1400	500	1900
Cadre moyen	1800	1300	3100
Employés	1500	3200	4700
Ouvriers	4500	2300	6800
Marge de B	10400	7900	18300

- 5° Dans la case marquée #, on écrit le nombre total de personnes sur lesquelles porte l'étude soit 18300.
- 6° On aurait pu étudier leur tranche d'âge (caractère quantitatif continu), leurs loisirs (caractère qualitatif nominal), leur région de naissance (caractère qualitatif que l'on peut ordonner suivant le numéro), leurs revenus (caractère quantitatif continu), leur pointure de chaussures (caractère quantitatif discret).
Mais pour les loisirs, il faut faire très attention, on risque de ne pas faire de partition (voir exercice de la page précédente sur l'équipement culturel des communes).

Exercice 2

Reconstituez un tableau à double entrée à partir des renseignements suivants :

Nombre de bovins mâles lourds : 25 000.

Nombre de bovins femelles lourdes : 9% de l'effectif total.

Nombre de bovins femelles de poids moyen : 22 000.

Nombre de bovins mâles de poids moyen : 18 000.

Nombre de bovins femelles de poids léger : 60 000.

Nombre de bovins mâles de poids léger : 45 000.

Exercice 3

Répartition des étudiants selon leur cycle d'étude et leur discipline.

Compléter le tableau suivant à l'aide des indications suivantes :

- Il n'y a qu'un cycle d'étude en IUT.
- Il y a 8 fois plus d'étudiants en 1er cycle qu'en troisième cycle.
- En 3^{ème} cycle, les étudiants en Droit ne représentent que 20% des effectifs.
- En Sciences, 60% des étudiants sont en premier cycle.
- Il y a autant d'étudiants en 1er cycle de Droit que d'étudiants en 2^{ème} et 3^{ème} cycle de Sciences.

	Droit	Lettres	Sciences	IUT	Total	Fréquences marginales
1 ^{er} cycle			1230	710		
2 ^{ème} cycle		108			855	
3 ^{ème} cycle						
Total	1108				4500	

Exercice 4 : Trouver l'erreur (extrait d'un sujet de baccalauréat série B).

Voici les résultats d'un examen ; malheureusement certains résultats sont effacés et il semble qu'il manque **le premier chiffre** d'un nombre inscrit dans une case. Essayez de reconstituer le tableau.

Résultats	Reçus	Collés	Totaux	Fréquence marginale
Série				
A		938		0,28
B		1036	2400	
C				0,15
D	978			
Totaux			9600	
Fréquence marginale	0,65			