

Croisement de deux variables catégorielles

1 Proportions et pourcentages (rappels)

Définition : proportion

On considère une population A d'effectif a et une sous population B de A d'effectif b .

La **proportion** de B par rapport à A est le nombre réel $\frac{b}{a}$.

Remarques

On parle aussi de fréquence.

Une proportion peut s'exprimer sous la forme d'une fraction ou d'un pourcentage.

Exemple

Jean-Kevin étudie une population de 120 oiseaux. Parmi ces derniers, 30 sont des mâles.

La proportion de mâle est de $\frac{30}{120}$ soit $\frac{1}{4}$ après simplification. On peut également exprimer cette proportion en pourcentage : $\frac{30}{120} \times 100 = 25$. Autrement dit, 25% des oiseaux étudiés par Jean-Kevin sont des mâles.

Propriété

Soient A, B et C trois ensembles tels que A soit inclus dans B et B inclus dans C. On notera $A \subset B \subset C$.

On note p_1 la proportion de A dans B et p_2 la proportion de B dans C.

La proportion de A dans C est $p_1 \times p_2$.

Exemple

On reprend les données de l'exemple précédent. Parmi les 30 mâles, 20% sont des mésanges.

La proportion de mésanges parmi tous les oiseaux observés est de $25\% \times 20\% = \frac{25}{100} \times \frac{20}{100} = \frac{1}{20} = 5\%$.

Parmi tous les oiseaux observés par Jean-Kevin, 5% sont des mésanges mâles.

Pour connaître le nombre de mésanges mâles : $120 \times \frac{1}{20} = 6$. Il a donc observé 6 mésanges mâles.

2 Tableaux croisés

Définition : tableau croisé d'effectifs

On considère une population A. On s'intéresse à deux caractères de cette population.

Lorsque l'on regroupe toutes ces informations dans un même tableau, on parle de **tableau croisé d'effectifs** ou bien de **tableau à deux entrées**.

Exemple

On s'intéresse aux élèves inscrits à l'UNSS d'un collège. On note cette population A. On s'intéresse particulièrement aux caractères suivants :

- le type de sport pratiqué (futsal, volley-ball, step, gymnastique)
- le sexe des élèves (féminin, masculin)

	Futsal : FS	Volley-Ball : V	Step : S	Gymnastique : G	Total
Sexe féminin : F	25	125	105	5	260
Sexe masculin : M	15	90	130	5	240
Total	40	215	235	10	500

Définition : Cardinal

On considère une population A. On considère un caractère F de cette population.

On appelle **cardinal** de F le nombre d'individus ayant le caractère F dans la population A.

On note ce nombre $Card(F)$.

Exemple

Dans notre exemple précédent, $Card(FS) = 25$ et $Card(M) = 240$.

Définition : Fréquence marginale

La **fréquence marginale** est une fréquence d'une sous population au sein d'une population.

Exemple

Dans l'exemple précédent, on souhaite calculer la fréquence marginale f de F parmi tous les élèves inscrits.

$$f = \frac{Card(F)}{Card(A)} = \frac{25}{500} = \frac{1}{20} = 0,05.$$

Définition : Fréquence conditionnelle

On considère deux variables A et B étudiées sur un même ensemble E.

On appelle **fréquence conditionnelle** de B dans A le nombre noté $f_A(B)$ défini par :

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

Exemple

On souhaite connaître la fréquence conditionnelle de S dans la variable F.

Cela signifie que l'on souhaite connaître la fréquence d'élèves pratiquant le step parmi les filles.

$$f_F(S) = \frac{\text{Card}(F \cap S)}{\text{Card}(F)} = \frac{105}{260} \approx 0,40$$

Il y a donc environ 40% des élèves qui font du step parmi les filles de l'UNSS.