

## DISTRIBUTION À DEUX CARACTÈRES

Une **distribution à deux caractères** correspond à l'ensemble des couples de données recueillies au cours d'une étude statistique portant sur deux caractères issus d'une même situation.

Dans une étude statistique, on donne le nom de **variable statistique** à tout caractère dont les données peuvent être différentes.

## CORRÉLATION

Étudier la **corrélation** entre deux variables statistiques, c'est décrire le **lien** entre deux caractères quantitatifs d'une distribution. Il est possible de qualifier le **type**, le **sens** et l'**intensité** d'une corrélation entre deux variables.

- Le type de corrélation correspond au **modèle mathématique** qui décrit le mieux le lien entre les variables.
- Une corrélation est dite **positive** ou **négative** selon le sens de variation des variables.

**Positive** Lorsque les valeurs d'une variable augmentent (ou diminuent), les valeurs de l'autre variable augmentent (ou diminuent) aussi.

**Négative** Lorsque les valeurs d'une variable augmentent (ou diminuent), les valeurs de l'autre variable diminuent (ou augmentent).

- Une corrélation est dite **nulle**, **faible**, **moyenne**, **forte** ou **parfaite** selon l'intensité du lien entre les variables.

## TABLEAU À DOUBLE ENTRÉE

Un **tableau à double entrée** permet de représenter une distribution à deux variables, et de qualifier le type et le sens de la corrélation qui peut exister entre les deux variables.

Dans un tableau à double entrée :

- l'une des variables est associée à la première colonne et l'autre variable, à la première ligne ;
- les données peuvent être regroupées en classes ;
- la ligne et la colonne des totaux indiquent les effectifs des deux variables ;
- la case inférieure droite comporte l'effectif total de la distribution ;
- la corrélation est dite linéaire lorsque la majorité des effectifs suivent l'une des deux diagonales.

Ex.: Chacun des couples suivants indique l'âge et le nombre de branches de 20 arbres de la même essence.

(12, 160) (13, 145) (13, 220) (15, 240) (16, 205) (19, 285) (21, 390) (22, 410)  
 (22, 350) (23, 360) (23, 450) (24, 325) (25, 480) (25, 455) (27, 320) (27, 475)  
 (27, 410) (29, 410) (30, 555) (33, 530)

Arbres de la même essence

Âge (années) \ Nombre de branches	Nombre de branches					Total
	[100, 200[	[200, 300[	[300, 400[	[400, 500[	[500, 600[	
[10, 15[	2	1	0	0	0	3
[15, 20[	0	3	0	0	0	3
[20, 25[	0	0	4	2	0	6
[25, 30[	0	0	1	5	0	6
[30, 35[	0	0	0	0	2	2
Total	2	4	5	7	2	20

Ce tableau à double entrée montre que la corrélation entre l'âge et le nombre de branches est linéaire et positive.

Certaines conclusions peuvent être tirées à partir du tableau. Voici quelques exemples :

- $\frac{2}{5}$  des arbres de cette distribution sont âgés de 25 ans ou plus ;
- 80 % des arbres de [300, 400[ branches ont moins de 25 ans ;
- $\frac{5}{6}$  des arbres âgés de [25, 30[ ans ont au moins 400 branches ;
- 45 % des arbres possèdent [200, 400[ branches.

## NUAGE DE POINTS

Un nuage de points permet de représenter une distribution à deux variables et de qualifier le type, le sens et l'intensité de la corrélation qui peut exister entre les deux variables.

Dans un nuage de points :

- l'une des variables est associée à l'axe des abscisses et l'autre variable, à l'axe des ordonnées ;
- chacun des couples de la distribution est représenté par un point ;
- la corrélation est dite linéaire lorsque les points tendent à former une droite oblique. Cette corrélation est dite nulle si les points sont distribués au hasard, et de plus en plus forte au fur et à mesure que les points se rapprochent d'une droite oblique. Voici les principales caractéristiques qu'on attribue à une corrélation linéaire :

**Corrélation nulle**

