



# Statistique

Représentation d'une distribution à  
deux caractères

# Représentation d'une distribution à deux caractères

## Distribution à deux caractères

1

2

Une étude statistique peut porter sur une **population** ou un **échantillon**.

- La **population** est l'ensemble des éléments visés par une recherche (individus, objet, jours, dates, etc.).
- L'**échantillon** est un sous-ensemble de la population, souvent plus pratique pour effectuer l'étude statistique.

# ○ Représentation d'une distribution à deux caractères

## Distribution à deux caractères

1

Dans une étude statistique, on nomme **variables statistiques** les caractéristiques étudiées dans une **population** ou un **échantillon**.

Ces variables sont soit qualitatives (couleur des yeux, moyen de transport, niveau de satisfaction, etc.) ou quantitatives (salaire, distance parcouru, nombre d'enfants, etc.).

2

# Représentation d'une distribution à deux caractères

## Distribution à deux caractères

1

2

Les variables quantitatives peuvent être discrètes (nombre d'enfants, nombre de jours de travail par semaine, etc.) ou continues (salaire annuelle, distance parcourue, la taille, etc.).

Généralement, la variable quantitative discrète est associée à un plus petit nombre de valeurs possibles, contrairement à la variable quantitative continue qui peut prendre un grand nombre de valeurs.

# Représentation d'une distribution à deux caractères

## Distribution à deux caractères

1

2

La **distribution à deux caractères** est l'ensemble des couples de données  $(x, y)$  recueillies au cours d'une étude statistique portant sur **deux caractères (variables) distincts** issues d'une même situation.

Exemple:  $x$  : le nombre de parties jouées  
 $y$  : le nombre de buts marqués

# Représentation d'une distribution à deux caractères

## Exemple 1

On considère les résultats en sciences et en mathématiques de 14 élèves à la fin de la première étape.

(72 , 80) (57 , 65) (91 , 77) (61 , 69) (73 , 71)  
(88 , 80) (73 , 75) (84 , 92) (66 , 76) (85 , 79)  
(93 , 82)(58 , 61)(74 , 77)(87 , 69)

L'ensemble des couples (résultats en sciences, résultats en maths) forme une distribution à deux caractères.

# Représentation d'une distribution à deux caractères

## But

1

2

L'objectif est d'analyser les caractéristiques d'une distribution à deux caractères, quantitatifs ou qualitatifs, pour ainsi déterminer s'il existe un lien entre eux. Nous verrons deux façon de représenter la distribution:

- Nuage de points
- Tableau à double entrée

# Représentation d'une distribution à deux caractères

## Nuage de points

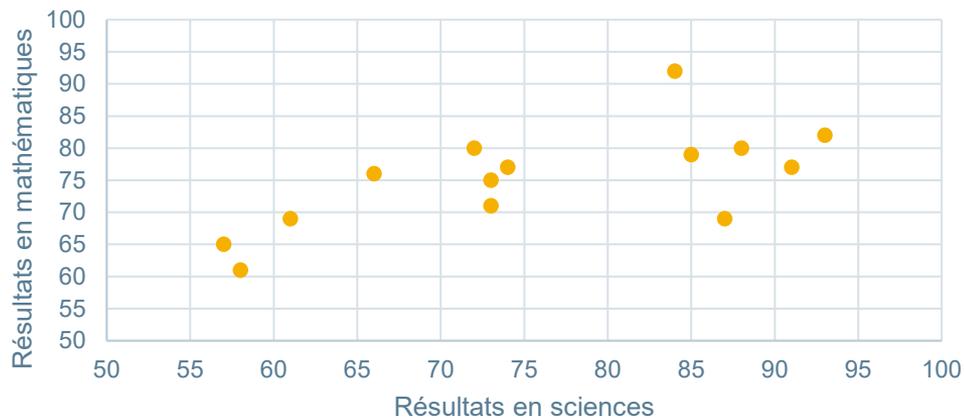
1

Le nuage de points est une représentation graphique qui permet de mettre en relation une distribution à 2 caractères quantitatifs.

Chaque point du nuage de points correspond à un couple  $(x, y)$  composé d'une valeur du caractère  $x$  et d'une valeur du caractère  $y$  de la distribution.

2

Résultats en sciences et en mathématiques de 14 élèves



# Représentation d'une distribution à deux caractères

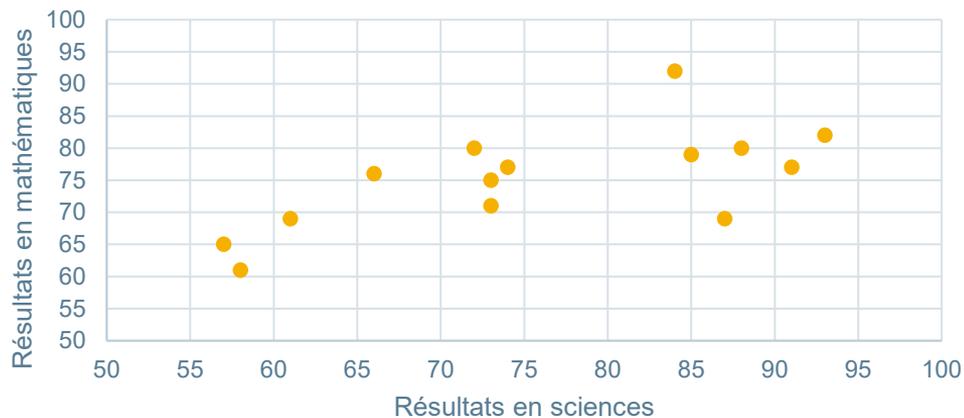
## Nuage de points

1

S'il existe un lien de dépendance entre les variables, nous mettrons la variable dépendante sur l'axe des ordonnées ( $y$ ) et l'indépendante sur l'axe des abscisses ( $x$ )

2

Résultats en sciences et en mathématiques de 14 élèves



# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1

Représentation d'une distribution à deux variables (quantitatives et/ou qualitatives). Chaque couple de données est compilé selon une valeur à l'horizontale et une valeur à la verticale.

2

### Exemple

Répartition des postes occupés dans une compagnie en fonction du niveau de scolarité				
Poste occupé	Niveau de scolarité			Total
	secondaire	collégial	universitaire	
directeur	2	2	10	14
technicien	4	22	8	34
opérateur	35	30	8	73
<b>Total</b>	<b>41</b>	<b>54</b>	<b>26</b>	<b>121</b>

# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1

Reprenons les résultats en sciences et en mathématiques de 14 élèves :

2

(72, 80) (57, 65) (91, 77) (61, 69) (73, 71)  
(88, 80) (73, 75) (84, 92) (66, 76) (85, 79)  
(93, 82)(58, 61)(74, 77)(87, 69)

Si les données recueillis pour une variable impliquent plus de 10 valeurs différentes pour celles-ci, nous allons regrouper ces valeurs en **classes**.

Il existe des conventions pour choisir un nombre de classes en fonction du nombre de valeurs différentes pour une variable. Dans le cadre du cours, nous allons choisir simplement de regrouper les données en 5 à 8 classes.

# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1  
2 Pour construire les classes, nous devons d'abord déterminer l'**étendue**.

$$\text{Étendue} = \text{valeur max} - \text{valeur min}$$

$$\text{Étendue (sciences)} = 93 - 57 = 36$$

$$\text{Étendue (mathématiques)} = 92 - 61 = 31$$

# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1

2

Ensuite, nous allons diviser l'étendue par le nombre de classes désirées pour obtenir l'**amplitude** (la largeur) de chaque classe. Puisque nous avons à peine plus de 10 valeurs différentes pour chaque variable, 5 classes seront suffisantes :

$$\textit{Amplitude (sciences)} = \frac{36}{5} = 7,2$$

$$\textit{Amplitude (mathématiques)} = \frac{31}{5} = 6,2$$

# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1

2

Afin de s'assurer d'inclure les plus petites et plus grandes valeurs dans nos classes, nous augmenterons légèrement l'amplitude obtenue pour s'assurer qu'elles y retrouveront.

Donc, on obtient une amplitude de 8 pour les classes des résultats en sciences et une amplitude de 7 pour les classes des résultats en mathématiques

On notera alors  $[57; 65[$  notre première classe pour sciences. Notez bien que le 65 est exclu de cette classe.

# Représentation d'une distribution à deux caractères

## Tableau à double entrée

1

2

Nous pouvons maintenant construire le tableau double entrée associé à la distribution des résultats en sciences et en mathématiques pour les 14 élèves:

Répartition de 14 élèves en fonction des résultats en sciences et des résultats en mathématiques						
Résultats en mathématiques	Résultats en sciences					Total
	[57;65[	[65;73[	[73;81[	[81;89[	[89;97[	
[61;68[	2	0	0	0	0	2
[68;75[	1	0	1	1	0	3
[75;82[	0	2	2	2	1	7
[82;89[	0	0	0	0	1	1
[89;96[	0	0	0	1	0	1
<b>Total</b>	3	2	3	4	2	14

# Représentation d'une distribution à deux caractères

Tableau à double  
entrée et nuage de  
points

1

2

Nuage de points	Tableau à double entrée
Données quantitatives discrètes ou continues  Généralement pour un plus petit nombre de données	Données qualitatives, quantitatives discrètes ou quantitatives continues  Grand nombre de données  Données regroupées en classe ou en intervalle (au besoin)



# Statistique

Représentation d'une distribution à  
deux caractères