

# I Définition - Nuage de points

SAVOIR :

→ Une série statistique est "double" (ou série "à 2 variables") quand on observe simultanément **deux** caractères  $x$  et  $y$  sur une population.

## Exemple

Par exemple, sur une population de pièces en acier, on mesure  $x$ , la teneur en carbone et  $y$  la charge nécessaire à la rupture de la pièce.

| Numéro pièce                       | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|------------------------------------|----|----|----|----|----|----|----|----|----|----|
| Teneur en carbone $x_i$ (en 0.01%) | 70 | 60 | 68 | 64 | 66 | 64 | 62 | 70 | 74 | 62 |
| Charge de rupture $y_i$ (en kg)    | 87 | 71 | 79 | 74 | 79 | 80 | 75 | 86 | 95 | 70 |

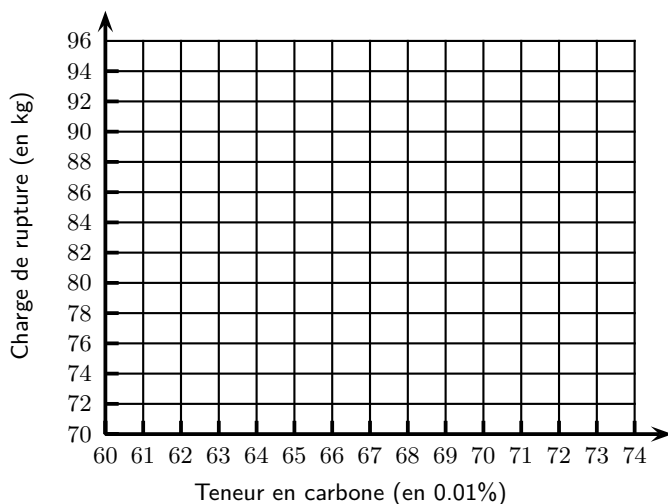
Il y a 10 couples de valeurs. L'effectif total est  $N = 10$ .

SAVOIR :

→ Le nuage de points est l'ensemble des  $N$  points dont les coordonnées sont les  $(x_i; y_i)$ .

## Exemple

Construire le nuage de points correspondant à l'exemple précédent.



# II Point moyen

SAVOIR :

→ Le point moyen d'un nuage de  $N$  points est le point  $G$  de coordonnées

$$\begin{cases} \bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} \\ \bar{y} = \frac{y_1 + y_2 + \dots + y_N}{N} \end{cases}$$

## Exemple

Déterminer les coordonnées du point moyen de l'exemple et le placer dans le graphique précédent.

$\bar{x} = \dots$

Et aussi à la calculatrice :

T.I.

[STAT] [EDIT]

Saisir les valeurs  $x_i$  et  $y_i$  dans L1 et L2

[STAT] [CALC] [2-VAR STATS]

Taper ensuite [L1] , [L2] ENTER

Lire alors  $n$  (effectif total) et  $\bar{x}$ , puis plus bas  $\bar{y}$ .

$\bar{y} = \dots$

Casio

[MENU] [STAT]

Saisir les valeurs  $x_i$  et  $y_i$  dans L1 et L2

[CALC] , puis régler dans [SET] :

2-VAR Xlist : L1, 2-VAR Ylist : L2 et 2-VAR Freq : 1

[EXE] puis [2VAR]

Lire alors  $n$  (effectif total) et  $\bar{x}$ , puis plus bas  $\bar{y}$ .

### III Ajustement affine

L'idée générale des statistiques à 2 variables est de rechercher s'il existe une relation entre les 2 caractères  $x$  et  $y$ . **Lorsque les points du nuage d'une série statistique sont approximativement alignés**, on cherche une fonction affine qui exprime de façon approchée  $y$  en fonction de  $x$ .

SAVOIR :

→ Un ajustement affine est justifié lorsque les points du nuage sont approximativement alignés.

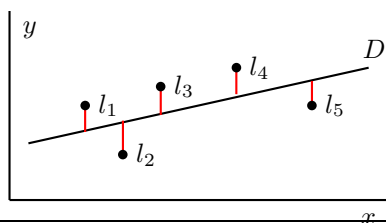
→ La droite d'équation  $y = ax + b$  que l'on trouve s'appelle la "**droite d'ajustement**".

→ Il existe diverses méthodes pour trouver une **droite d'ajustement**, nous en voyons 2 cette année :

#### III.1 Méthode graphique

On trace **au jugé** une droite en s'efforçant d'équilibrer le nombre de points situés de part et d'autre. Cette méthode est peu précise.

#### III.2 Méthode des moindres carrés



La droite d'ajustement donnée par cette méthode est celle qui rend minimale la somme des carrés de toutes les longueurs sur la figure ( $l_1^2 + l_2^2 + \dots + l_n^2$ ).

SAVOIR :

→ Cette droite est appelée **droite de régression de  $y$  en  $x$** .

→ Les coefficients  $a$  et  $b$  de son équation  $y = ax + b$  sont calculés directement à la calculatrice (voir ci-dessous).

→ Pour représenter la droite dans le repère, on doit placer 2 points. Un premier point utilisable est le point moyen, qui est toujours dessus, et un autre point peut être trouvé en faisant un tableau de valeurs à la calculatrice.

Texas Instrument

**STAT**, **EDIT**

Saisir les valeurs  $x_i$  et  $y_i$  L1 et L2

**STAT**, **CALC**, **LinReg(ax+b)**

Taper ensuite **L1, L2 ENTER**

Lire alors  $a$  et  $b$ .

Casio

**MENU**, **STAT**

Saisir les valeurs  $x_i$  et  $y_i$  dans L1 et L2

**CALC**, puis régler dans **SET** :

2-VAR Xlist : L1, 2-VAR Ylist : L2 et 2-VAR Freq : 1

**EXE** puis **REG** **X** (puis **ax + b** pour certaines calculatrices) .

Lire alors  $a$  et  $b$ .

#### Exemple

1°) Reprendre la série du I. Déterminer à la calculatrice (*arrondir à 0,1 près*) une équation de la droite de régression de  $y$  en  $x$  par la méthode des moindres carrés.

2°) Représenter cette droite dans le repère.

#### III.3 Utilisation des ajustements : estimations et prévisions

SAVOIR :

→ On utilise l'équation de la droite pour donner une **estimation de couples de valeurs** ne figurant pas dans la série statistique de départ.

#### Exemple

1°) En utilisant la droite de régression (trouvée à l'exemple précédent), **ESTIMER** la charge à laquelle peut résister une pièce dont la teneur en carbone est 0,72%.

2°) De la même manière, **PREVOIR** la teneur en carbone d'une pièce qui résisterait à 100 kg de charge.