

## 2 Séries statistiques à deux variables

On étudie simultanément deux caractères d'une population.

La série statistique à deux variables est donnée par des couples de valeur  $(x_i ; y_i)$ .

On la représente graphiquement par un **nuage de points**.

### 1. Point moyen

On appelle **point moyen** d'un nuage de  $n$  points  $M_i(x_i ; y_i)$  le point  $G$  de coordonnées :

$$x_G = \bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) ; y_G = \bar{y} = \frac{1}{n} (y_1 + y_2 + \dots + y_n).$$

### 2. Ajustement affine

Lorsque le nuage a une forme allongée, il est possible de tracer une droite  $D$  au voisinage de ses points.

On dit alors que l'on a un **ajustement affine**.

### 3. Ajustement affine par la méthode des moindres carrés

Par la méthode des moindres carrés, on détermine deux droites, appelées **droites de régression**.

La droite de régression de  $y$  en  $x$  a pour équation  $y = ax + b$ .

La droite de régression de  $x$  en  $y$  a pour équation  $x = a'y + b'$ .

Ces deux droites passent par le **point moyen G** du nuage de points.

Les équations de ces droites sont obtenues avec une calculatrice ou avec un logiciel (voir *Fiche méthode 27*).

Le **coefficient de corrélation  $r$**  donné par la calculatrice permet de comparer la qualité de deux ajustements. Plus ce coefficient est **proche de 1 ou - 1**, meilleur est l'ajustement.

Les droites de régression permettent d'obtenir **des estimations** :

– pour obtenir une estimation de  **$y$**  pour une valeur de  **$x$**  donnée, on utilise l'équation de la droite de régression **de  $y$  en  $x$** .

– pour obtenir une estimation **de  $x$**  pour une valeur de  **$y$**  donnée, on utilise l'équation de la droite de régression  **$x$  en  $y$** .