

Définitions

Soit $n \in \mathbb{N}^*$ et soient $x = (x_1; \dots; x_n)$ et $y = (y_1; \dots; y_n)$ deux séries statistiques à n termes réels.

- 1) L'ensemble des points $M_1(x_1; y_1); \dots; M_n(x_n; y_n)$ est appelé le nuage de points associé à la série statistique double $(x; y)$.
- 2) Le point $\overline{M}(\bar{x}; \bar{y})$, de coordonnées les moyennes $\bar{x}$ et $\bar{y}$ séries statistiques x et y , est appelé le point moyen de la série statistique double $(x; y)$.

- 3) Le nombre réel

$$\text{cov}(x; y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

est appelé la covariance des séries statistiques x et y .

- 4) La droite d'équation $y = ux + v$ avec

$$\begin{cases} u &= \frac{\text{cov}(x; y)}{\text{var}(x)} \\ v &= \bar{y} - u\bar{x} \end{cases}$$

est appelée la droite de régression (ou d'ajustement) des séries statistiques x et y par la méthode des moindres carrés.

- 5) Le nombre réel

$$r = \frac{\text{cov}(x; y)}{\sigma_x \sigma_y}$$

où σ_x et σ_y sont les écarts-types des séries statistiques x et y , est appelé le coefficient de corrélation linéaire des séries statistiques x et y .