

Théorème

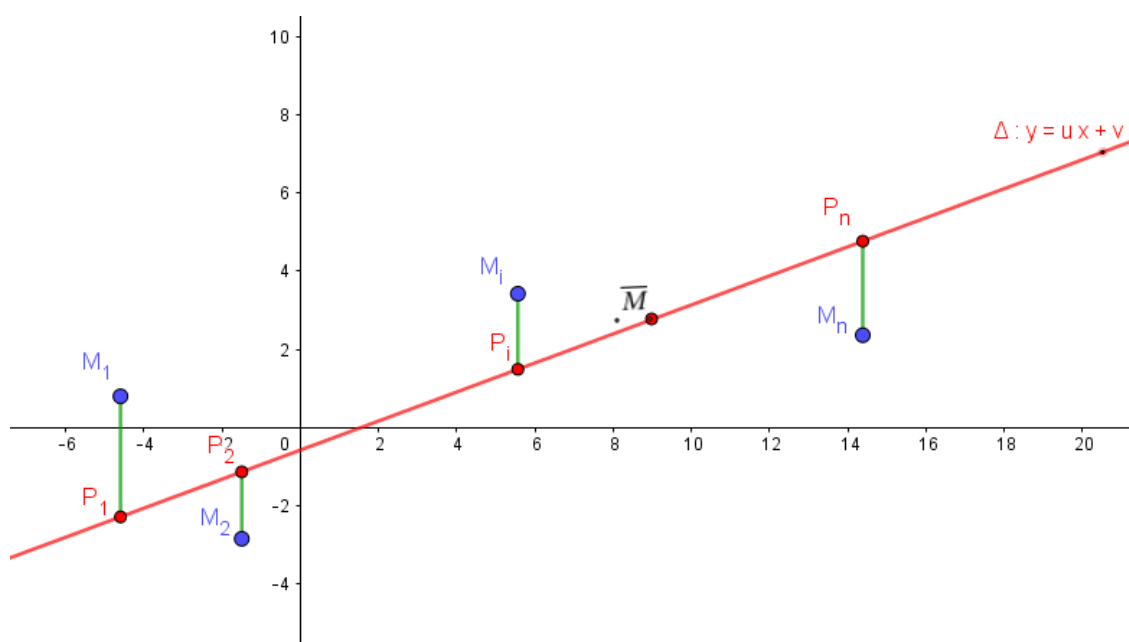
Soit $n \in \mathbb{N}^*$ et soient $x = (x_1; \dots; x_n)$ et $y = (y_1; \dots; y_n)$ deux séries statistiques à n termes réels.

Soient $M_1(x_1; y_1); \dots; M_n(x_n; y_n)$ les points de coordonnées respectives $(x_1; y_1); \dots; M(x_n; y_n)$ dans un repère orthogonal du plan.

Soit $\overline{M}(\bar{x}; \bar{y})$ le point de coordonnées les moyennes $\bar{x}$ et $\bar{y}$ séries statistiques x et y .

Soit $(u; v)$ un couple de nombres réels et soit Δ la droite d'équation $y = ux + v$.

Pour tout entier i entre 1 et n , soit P_i le point d'abscisse x_i de la droite Δ .



- 1) Il existe un unique couple d'entiers $(u; v)$ tel que la somme $P_1 M_1^2 + \dots + P_n M_n^2 = \sum_{i=1}^n P_i M_i^2$ soit minimale.

$$\begin{cases} u = \frac{\text{cov}(x; y)}{\text{var}(x)} \\ v = \bar{y} - u\bar{x} \end{cases} \quad \text{avec} \quad \text{cov}(x; y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

- 2) Soit $r = \frac{\text{cov}(x; y)}{\sigma_x \sigma_y}$ où σ_x et σ_y sont les écarts-types des séries statistiques x et y .

a) On a $-1 \leq r \leq 1$.

b) $r = \pm 1$ si et seulement si les points $P_1, \dots, P_n$ sont alignés.

c) Les points $M_1(x_1; y_1); \dots; M_n(x_n; y_n)$ sont d'autant plus proches de la droite Δ que r est proche de ± 1 .