

06-02 Ajustement affine

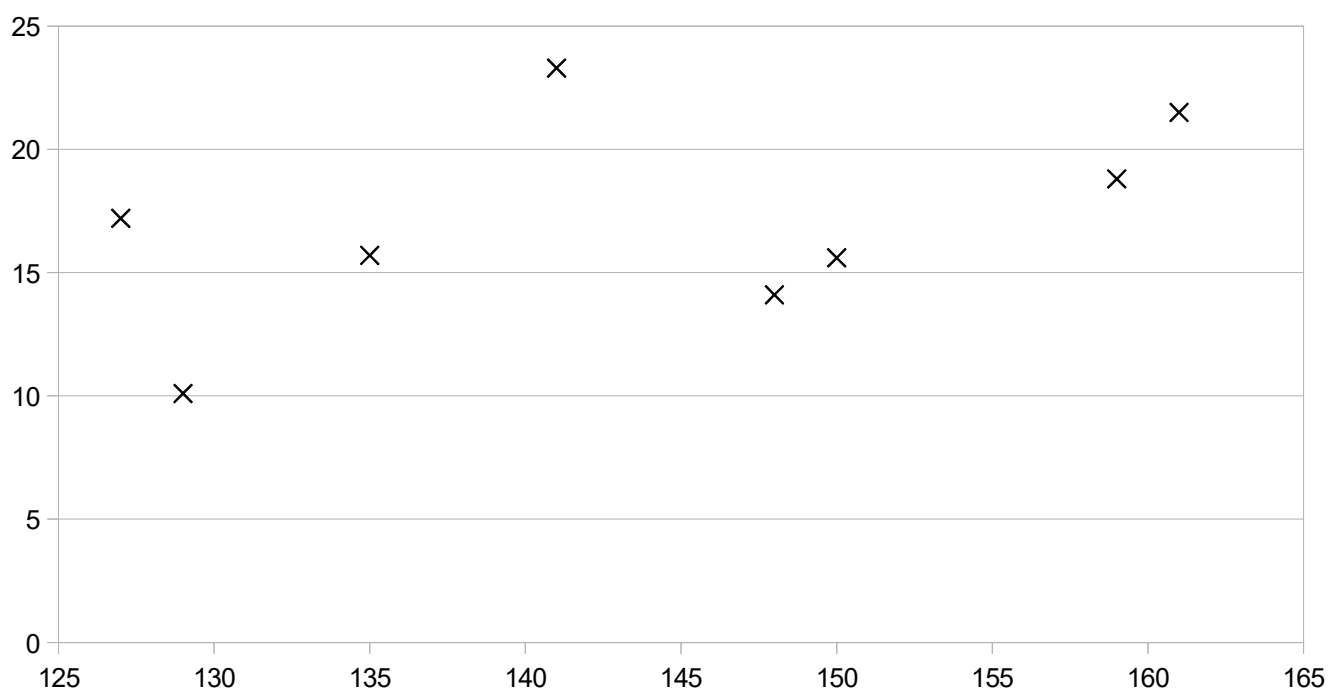
Définitions

On appelle **droite de régression** d'un nuage de points une droite qui « approche » les points du nuage. La recherche d'une telle droite est appelée **ajustement affine** d'un nuage de points.

Remarques

- Un nuage de points peut admettre plusieurs droites de régression différentes selon la méthode utilisée.
- Quel que soit l'ajustement affine, le point moyen du nuage doit appartenir à la droite de régression.

Exemple



La droite de régression ci-dessus, tracée « au jugé », a pour équation :

Une ville ayant 200 jours de soleil par an aura donc une température moyenne proche de

06-03 Variance et covariance

Définitions et notation

La **variance** d'une série de valeurs $x_1 ; x_2 ; \dots ; x_n$ est la moyenne du carré des écarts à la moyenne. L'**écart-type** de la série est la racine carrée de la variance. On le note $\sigma(x) = \sqrt{V(x)}$.

Remarques

- La variance et l'écart-type sont des **indicateurs de dispersion** qui évaluent l'hétérogénéité d'une série.
- On peut calculer la variance grâce à la formule **Variance = Moyenne des carrés – Carré de la moyenne**.

Exemple

On considère les séries ci-dessous.

x_i	161	141	127	150	159	129	148	135
y_i	21,5	23,3	17,2	15,6	18,8	10,1	14,1	15,7

On a : $Vx \approx 146$; $Vy \approx 15$; $\sigma x \approx 12$; $\sigma y \approx 4$.

Définitions et notation

La **covariance** de deux séries de valeurs $(x_1 ; x_2 ; \dots ; x_n)$ et $(y_1 ; y_2 ; \dots ; y_n)$ est la moyenne des produits terme à terme des écarts à la moyenne.

On la note $\text{cov}(x ; y) = \frac{1}{n} [(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})]$

Remarques

- La covariance est un **indicateur de dispersion** qui s'éloigne de 0 quand les séries sont interdépendantes.
- On peut calculer la covariance grâce à : **Covariance = Moyenne des produits – Produit des moyennes**.

$$\text{cov}(x ; y) = \frac{1}{n} (x_1 y_1 + x_2 y_2 + \dots + x_n y_n) - \bar{x} \bar{y}$$

Exemple

La covariance des deux séries précédentes vaut :

$$\frac{1}{8} (161 \times 21,5 + 141 \times 23,3 + \dots + 135 \times 15,7) - 143,75 \times 17,0375 \approx 22$$