

Régression Linéaire Simple – Méthode des Moindres Carrés

La régression linéaire simple est une méthode statistique permettant de modéliser la relation entre une variable explicative (X) et une variable à expliquer (Y) par une droite. L'objectif de la méthode des moindres carrés est de trouver la droite qui minimise la somme des carrés des erreurs (écarts entre valeurs observées et valeurs prédites).

Formules générales

Modèle : $y = b_0 + b_1 \cdot x$

Erreur : $e_i = y_i - (b_0 + b_1 \cdot x_i)$

Critère : Minimiser $S(b_0, b_1) = \text{Somme } (y_i - b_0 - b_1 \cdot x_i)^2$

Solutions analytiques :

$b_1 = \text{Somme}[(x_i - \text{moyenne}(X)) \cdot (y_i - \text{moyenne}(Y))] / \text{Somme}[(x_i - \text{moyenne}(X))^2]$

$b_0 = \text{moyenne}(Y) - b_1 \cdot \text{moyenne}(X)$

Exemple : Surface et Prix

On souhaite modéliser la relation entre la surface d'un appartement (X, en m²) et son prix (Y, en milliers d'euros). Voici les données :

Surfaces = [50, 60, 70, 80, 90, 100]

Prix = [150, 180, 200, 220, 240, 280]

Surface (m ²)	Prix observé (k€)
50	150
60	180
70	200
80	220
90	240
100	280

Calculs étape par étape

- Moyenne(X) = 75

- Moyenne(Y) = 211,67

- Somme $(x_i - \text{Moyenne}(X))^2 = 1750$

- Somme $(x_i - \text{Moyenne}(X))(y_i - \text{Moyenne}(Y)) = 4250$

Résultats :

$b_1 = 4250 / 1750 \approx 2,43$

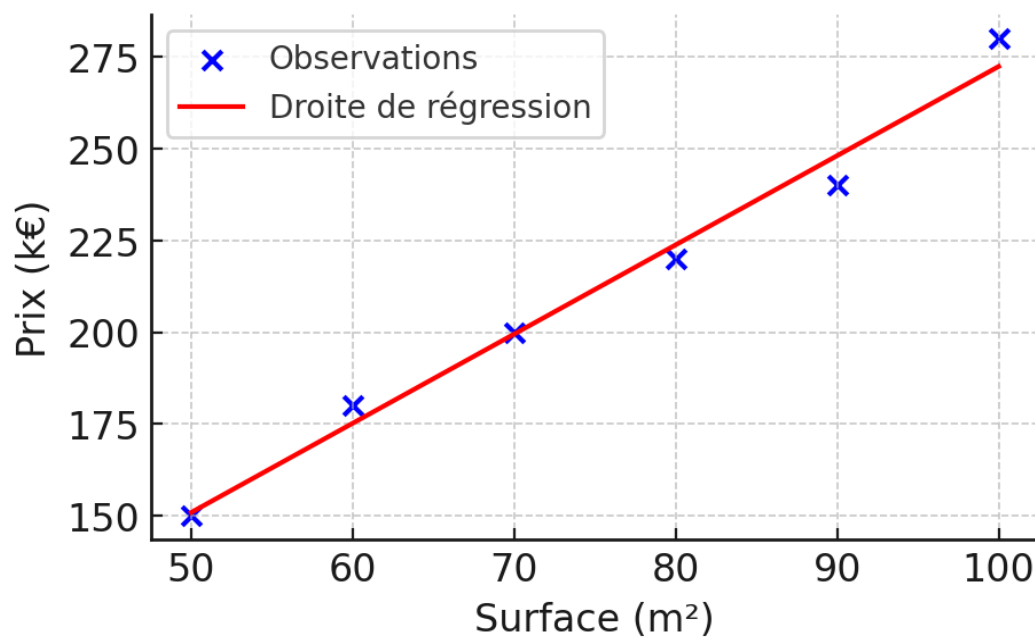
$b_0 = 211,67 - 2,43 \times 75 \approx 29,42$

Équation de la droite de régression

Prix $\approx 29,42 + 2,43 \times \text{Surface}$

Tableau des prédictions

Surface (m²)	Prix observé (k€)	Prix prédit (k€)
50	150	150.92
60	180	175.17
70	200	199.42
80	220	223.67
90	240	247.92
100	280	272.17



Conclusion

La méthode des moindres carrés fournit une droite de régression : $\text{Prix} \approx 29,42 + 2,43 \times \text{Surface}$.
Interprétation : en moyenne, chaque m² supplémentaire ajoute environ 2,43 k€ au prix.
Cette droite passe toujours par le point (moyenne(X), moyenne(Y)) = (75, 211,67).