

1 Calcul des expressions pour M et B par la méthode des moindres carrés

Imaginons trois points expérimentaux (x_1, y_1) , (x_2, y_2) et (x_3, y_3) qui ne sont pas parfaitement alignés dans le plan cartésien. Lorsqu'on cherche à obtenir une **droite de régression linéaire** d'équation $Y = MX + B$, on applique la **méthode des moindres carrés** afin de calculer les meilleures valeurs possibles de la **pente M** et de l'**ordonnée à l'origine B** afin que la droite obtenue passe le plus près possible de tous les points expérimentaux.

Plus précisément, on cherche à minimiser la somme des carrés des différences entre les valeurs des ordonnées prédites par la droite de régression $(Mx_i + B)$ et les valeurs expérimentales (y_i) . Dans le cas d'une droite de régression linéaire passant le plus près possible de trois points expérimentaux, la fonction f que nous cherchons à minimiser est :

$$f = (Mx_1 + B - y_1)^2 + (Mx_2 + B - y_2)^2 + (Mx_3 + B - y_3)^2$$

Dans cette fonction f , notons que x_1, x_2, x_3, y_1, y_2 et y_3 ne sont pas des variables, ce sont des valeurs expérimentales connues et fixes. C'est plutôt M et B qui sont les deux variables que l'on cherche à calculer, afin de minimiser la valeur de la fonction f .

Quand la fonction f atteint sa valeur minimale, les deux conditions suivantes sont respectées :

$$\frac{\partial f}{\partial M} = 0 \quad \text{et} \quad \frac{\partial f}{\partial B} = 0$$

En effectuant la dérivée partielle de f par rapport à M et en posant que celle-ci doit être égale à zéro, on obtient une première équation :

$$\begin{aligned} \frac{\partial}{\partial M} [(Mx_1 + B - y_1)^2 + (Mx_2 + B - y_2)^2 + (Mx_3 + B - y_3)^2] &= 0 \\ 2 \cdot (Mx_1 + B - y_1) \cdot x_1 + 2 \cdot (Mx_2 + B - y_2) \cdot x_2 + 2 \cdot (Mx_3 + B - y_3) \cdot x_3 &= 0 \\ (Mx_1 + B - y_1) \cdot x_1 + (Mx_2 + B - y_2) \cdot x_2 + (Mx_3 + B - y_3) \cdot x_3 &= 0 \\ (Mx_1^2 + Bx_1 - x_1y_1) + (Mx_2^2 + Bx_2 - x_2y_2) + (Mx_3^2 + Bx_3 - x_3y_3) &= 0 \\ M \cdot (x_1^2 + x_2^2 + x_3^2) + B \cdot (x_1 + x_2 + x_3) - (x_1y_1 + x_2y_2 + x_3y_3) &= 0 \end{aligned}$$

(Entre la deuxième et la troisième ligne, on a simplement divisé partout par 2 afin de simplifier.)

On peut généraliser ce résultat ainsi :

$$\boxed{M \cdot \sum x^2 + B \cdot \sum x - \sum xy = 0} \quad (1)$$

D'autre part, en effectuant la dérivée partielle de f par rapport à B et en posant que celle-ci doit être égale à zéro, on obtient une seconde équation :

$$\begin{aligned} \frac{\partial}{\partial B} [(Mx_1 + B - y_1)^2 + (Mx_2 + B - y_2)^2 + (Mx_3 + B - y_3)^2] &= 0 \\ 2 \cdot (Mx_1 + B - y_1) + 2 \cdot (Mx_2 + B - y_2) + 2 \cdot (Mx_3 + B - y_3) &= 0 \\ (Mx_1 + B - y_1) + (Mx_2 + B - y_2) + (Mx_3 + B - y_3) &= 0 \\ M \cdot (x_1 + x_2 + x_3) + (B + B + B) - (y_1 + y_2 + y_3) &= 0 \end{aligned}$$

(Entre la deuxième et la troisième ligne, on a simplement divisé partout par 2 afin de simplifier.)

On peut généraliser ce résultat ainsi :

$$\boxed{M \cdot \sum x + n \cdot B - \sum y = 0} \quad (2)$$

où n est le nombre de points (x, y) dans nos données expérimentales.

Ensemble, les équations (1) et (2) constituent un système de deux équations et deux inconnues : M et B .

En isolant B dans l'équation (2), on obtient :

$$B = \frac{\Sigma y - M \cdot \Sigma x}{n} \quad (3)$$

En substituant cette expression pour B dans l'équation (1) et en multipliant la ligne entière par n afin de faire disparaître la fraction, on obtient :

$$\begin{aligned} M \cdot \Sigma x^2 + \left[\frac{\Sigma y - M \cdot \Sigma x}{n} \right] \cdot \Sigma x - \Sigma xy &= 0 \\ n \cdot M \cdot \Sigma x^2 + [\Sigma y - M \cdot \Sigma x] \cdot \Sigma x - n \cdot \Sigma xy &= 0 \\ n \cdot M \cdot \Sigma x^2 + \Sigma x \cdot \Sigma y - M \cdot (\Sigma x)^2 - n \cdot \Sigma xy &= 0 \\ M \cdot (n \cdot \Sigma x^2 - (\Sigma x)^2) &= n \cdot \Sigma xy - \Sigma x \cdot \Sigma y \\ M &= \frac{n \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{n \cdot \Sigma x^2 - (\Sigma x)^2} \end{aligned}$$

On obtient donc l'équation générale pour la pente M de la droite de régression linéaire :

$$M = \frac{n \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{n \cdot \Sigma x^2 - (\Sigma x)^2} \quad (4)$$

En remplaçant cette expression pour M dans l'équation (3) et en multipliant la ligne entière par n afin de faire disparaître la fraction, on obtient :

$$\begin{aligned} B &= \frac{\Sigma y - \left[\frac{n \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{n \cdot \Sigma x^2 - (\Sigma x)^2} \right] \cdot \Sigma x}{n} \\ n \cdot B &= \Sigma y - \left[\frac{n \cdot \Sigma xy - \Sigma x \cdot \Sigma y}{n \cdot \Sigma x^2 - (\Sigma x)^2} \right] \cdot \Sigma x \\ n \cdot B &= \frac{\Sigma y \cdot (n \cdot \Sigma x^2 - (\Sigma x)^2)}{n \cdot \Sigma x^2 - (\Sigma x)^2} - \frac{\Sigma x \cdot (n \cdot \Sigma xy - \Sigma x \cdot \Sigma y)}{n \cdot \Sigma x^2 - (\Sigma x)^2} \\ n \cdot B &= \frac{n \cdot \Sigma x^2 \cdot \Sigma y - (\Sigma x)^2 \cdot \Sigma y - n \cdot \Sigma x \cdot \Sigma xy + (\Sigma x)^2 \cdot \Sigma y}{n \cdot \Sigma x^2 - (\Sigma x)^2} \\ n \cdot B &= \frac{n \cdot \Sigma x^2 \cdot \Sigma y - n \cdot \Sigma x \cdot \Sigma xy}{n \cdot \Sigma x^2 - (\Sigma x)^2} \\ B &= \frac{\Sigma x^2 \cdot \Sigma y - \Sigma x \cdot \Sigma xy}{n \cdot \Sigma x^2 - (\Sigma x)^2} \end{aligned}$$

(À la ligne 3, on a multiplié Σy par « 1 » afin d'exprimer les deux termes avec le même dénominateur commun.

Entre la ligne 4 et la ligne 5, le terme « $-(\Sigma x)^2 \cdot \Sigma y$ » s'annule avec le terme « $+(\Sigma x)^2 \cdot \Sigma y$ ».)

On obtient donc l'équation générale pour l'ordonnée à l'origine B de la droite de régression linéaire :

$$B = \frac{\Sigma x^2 \cdot \Sigma y - \Sigma x \cdot \Sigma xy}{n \cdot \Sigma x^2 - (\Sigma x)^2} \quad (5)$$

Pour calculer les dérivées partielles associées à la pente M obtenue à l'équation (4), posons d'abord que celle-ci est composée d'un numérateur μ et d'un dénominateur D :

$$\mu = n \cdot (\Sigma xy) - (\Sigma x) \cdot (\Sigma y) \quad \text{et} \quad D = n \cdot (\Sigma x^2) - (\Sigma x)^2$$

Pour le numérateur μ , effectuons d'abord un développement en supposant que nous avons seulement trois points expérimentaux (x_1, y_1) , (x_2, y_2) et (x_3, y_3) . Puis, en effectuant la dérivée partielle de μ par rapport à x_1 et de μ par rapport à y_1 , on peut facilement généraliser pour les dérivées partielles par rapport à x_i et à y_i .

$$\begin{aligned} \mu &= n \cdot (\Sigma xy) - (\Sigma x) \cdot (\Sigma y) \\ &= n \cdot (x_1 y_1 + x_2 y_2 + x_3 y_3) - (x_1 + x_2 + x_3) \cdot (y_1 + y_2 + y_3) \end{aligned}$$

La dérivée partielle de μ par rapport à x_1 est :

$$\frac{\partial \mu}{\partial x_1} = n y_1 - (y_1 + y_2 + y_3)$$

On peut généraliser ce résultat ainsi :

$$\boxed{\frac{\partial \mu}{\partial x_i} = n y_i - \Sigma y} \quad (6)$$

La dérivée partielle de μ par rapport à y_1 est :

$$\frac{\partial \mu}{\partial y_1} = n x_1 - (x_1 + x_2 + x_3)$$

On peut généraliser ce résultat ainsi :

$$\boxed{\frac{\partial \mu}{\partial y_i} = n x_i - \Sigma x} \quad (7)$$

Pour le numérateur D , effectuons d'abord un développement en supposant que nous avons seulement trois points expérimentaux (x_1, y_1) , (x_2, y_2) et (x_3, y_3) . Puis, en effectuant la dérivée partielle de D par rapport à x_1 et de D par rapport à y_1 , on peut facilement généraliser pour les dérivées partielles par rapport à x_i et à y_i .

$$\begin{aligned} D &= n \cdot (\Sigma x^2) - (\Sigma x)^2 \\ &= n \cdot (x_1^2 + x_2^2 + x_3^2) - (x_1 + x_2 + x_3)^2 \end{aligned}$$

La dérivée partielle de D par rapport à x_1 est :

$$\frac{\partial D}{\partial x_1} = 2n x_1 - 2 \cdot (x_1 + x_2 + x_3)$$

On peut généraliser ce résultat ainsi :

$$\boxed{\frac{\partial D}{\partial x_i} = 2n x_i - 2 \Sigma x} \quad (8)$$

Par ailleurs, puisque le dénominateur D ne dépend pas des paramètres y_i , on peut écrire directement :

$$\boxed{\frac{\partial D}{\partial y_i} = 0} \quad (9)$$

La pente M est un quotient de deux fonctions μ et D qu'on peut réexprimer comme un produit :

$$M = \frac{\mu}{D} = \mu D^{-1}$$

La dérivée partielle de M par rapport à x_i est :

$$\frac{\partial M}{\partial x_i} = \frac{1}{D} \frac{\partial \mu}{\partial x_i} - \frac{\mu}{D^2} \frac{\partial D}{\partial x_i}$$

En utilisant les équations (6) et (8) et en utilisant le fait que $\mu D^{-1} = M$, on obtient :

$$\begin{aligned} \frac{\partial M}{\partial x_i} &= \frac{1}{D} (ny_i - \Sigma y) - \frac{\mu}{D^2} (2nx_i - 2\Sigma x) \\ &= \frac{1}{D} (ny_i - \Sigma y) - \frac{M}{D} (2nx_i - 2\Sigma x) \\ &= \frac{ny_i}{D} - \frac{\Sigma y}{D} - \frac{2nMx_i}{D} + \frac{2M\Sigma x}{D} \end{aligned}$$

$$\boxed{\frac{\partial M}{\partial x_i} = \left[-\frac{2nM}{D} \right] x_i + \left[\frac{n}{D} \right] y_i + \left[\frac{2M\Sigma x - \Sigma y}{D} \right]} \quad (10)$$

La dérivée partielle de M par rapport à y_i est :

$$\frac{\partial M}{\partial y_i} = \frac{1}{D} \frac{\partial \mu}{\partial y_i} - \frac{\mu}{D^2} \frac{\partial D}{\partial y_i}$$

En utilisant les équations (7) et (9), on obtient :

$$\begin{aligned} \frac{\partial M}{\partial y_i} &= \frac{1}{D} (nx_i - \Sigma x) - \frac{\mu}{D^2} (0) \\ &= \frac{nx_i}{D} - \frac{\Sigma x}{D} \end{aligned}$$

$$\boxed{\frac{\partial M}{\partial y_i} = \left[\frac{n}{D} \right] x_i + \left[-\frac{\Sigma x}{D} \right]} \quad (11)$$

En utilisant les équations (10) et (11), nous pouvons désormais calculer les dérivées partielles de M par rapport à n'importe quel paramètre x_i et y_i . Ainsi, pour un graphique contenant n points (x, y) , l'erreur absolue δM associée à la pente de la droite de régression linéaire peut être calculée en appliquant la **méthode différentielle** pour la propagation des erreurs :

$$\delta M = \left| \frac{\partial M}{\partial x_1} \right| \delta x_1 + \left| \frac{\partial M}{\partial x_2} \right| \delta x_2 + \dots + \left| \frac{\partial M}{\partial x_n} \right| \delta x_n + \left| \frac{\partial M}{\partial y_1} \right| \delta y_1 + \left| \frac{\partial M}{\partial y_2} \right| \delta y_2 + \dots + \left| \frac{\partial M}{\partial y_n} \right| \delta y_n$$

Pour calculer les dérivées partielles associées à l'ordonnée à l'origine B obtenue à l'équation (5), posons d'abord que celle-ci est composée d'un numérateur β et d'un dénominateur D :

$$\beta = (\Sigma x^2) \cdot (\Sigma y) - (\Sigma x) \cdot (\Sigma xy) \quad \text{et} \quad D = n \cdot (\Sigma x^2) - (\Sigma x)^2$$

Pour le numérateur β , effectuons d'abord un développement en supposant que nous avons seulement trois points expérimentaux (x_1, y_1) , (x_2, y_2) et (x_3, y_3) . Puis, en effectuant la dérivée partielle de β par rapport à x_1 et de β par rapport à y_1 , on peut facilement généraliser pour les dérivées partielles par rapport à x_i et à y_i .

$$\begin{aligned} \beta &= (\Sigma x^2) \cdot (\Sigma y) - (\Sigma x) \cdot (\Sigma xy) \\ &= (x_1^2 + x_2^2 + x_3^2) \cdot (y_1 + y_2 + y_3) - (x_1 + x_2 + x_3) \cdot (x_1 y_1 + x_2 y_2 + x_3 y_3) \\ &= (x_1^2 + x_2^2 + x_3^2) \cdot (y_1 + y_2 + y_3) - x_1^2 y_1 - x_1 x_2 y_2 - x_1 x_3 y_3 - x_1 x_2 y_1 - x_2^2 y_2 - x_2 x_3 y_3 - x_1 x_3 y_1 - x_2 x_3 y_2 - x_3^2 y_3 \end{aligned}$$

La dérivée partielle de β par rapport à x_1 est :

$$\begin{aligned} \frac{\partial \beta}{\partial x_1} &= 2x_1 \cdot (y_1 + y_2 + y_3) - 2x_1 y_1 - x_2 y_2 - x_3 y_3 - x_2 y_1 - x_3 y_1 \\ &= 2x_1 \cdot (y_1 + y_2 + y_3) - x_1 y_1 - x_2 y_2 - x_3 y_3 - x_1 y_1 - x_2 y_1 - x_3 y_1 \\ &= 2x_1 \cdot (y_1 + y_2 + y_3) - (x_1 y_1 + x_2 y_2 + x_3 y_3) - y_1 \cdot (x_1 + x_2 + x_3) \end{aligned}$$

(Remarquez que le terme $-2x_1 y_1$ a été scindé en deux termes $-x_1 y_1$ afin de regrouper stratégiquement les termes dans l'équation.)

On peut généraliser ce résultat ainsi :

$$\boxed{\frac{\partial \beta}{\partial x_i} = 2x_i \Sigma y - \Sigma xy - y_i \Sigma x} \quad (12)$$

La dérivée partielle de β par rapport à y_1 est :

$$\begin{aligned} \frac{\partial \beta}{\partial y_1} &= (x_1^2 + x_2^2 + x_3^2) - x_1^2 - x_1 x_2 - x_1 x_3 \\ &= (x_1^2 + x_2^2 + x_3^2) - x_1 x_1 - x_1 x_2 - x_1 x_3 \\ &= (x_1^2 + x_2^2 + x_3^2) - x_1 (x_1 + x_2 + x_3) \end{aligned}$$

(Remarquez que le terme $-x_1^2$ a été réécrit comme $-x_1 x_1$ afin de regrouper stratégiquement les termes dans l'équation.)

On peut généraliser ce résultat ainsi :

$$\boxed{\frac{\partial \beta}{\partial y_i} = \Sigma x^2 - x_i \Sigma x} \quad (13)$$

Le dénominateur D dans l'équation de l'ordonnée à l'origine B est exactement le même que celui qui apparaît dans l'équation de la pente M . Les dérivées partielles de D ont déjà été obtenues aux équations (8) et (9).

L'ordonnée à l'origine B est un quotient de deux fonctions β et D qu'on peut réexprimer comme un produit :

$$B = \frac{\beta}{D} = \beta D^{-1}$$

La dérivée partielle de B par rapport à x_i est :

$$\frac{\partial B}{\partial x_i} = \frac{1}{D} \frac{\partial \beta}{\partial x_i} - \frac{\beta}{D^2} \frac{\partial D}{\partial x_i}$$

En utilisant les équations (12) et (8) et en utilisant le fait que $\beta D^{-1} = B$, on obtient :

$$\begin{aligned} \frac{\partial B}{\partial x_i} &= \frac{1}{D} (2x_i \Sigma y - \Sigma x y - y_i \Sigma x) - \frac{\beta}{D^2} (2n x_i - 2 \Sigma x) \\ &= \frac{1}{D} (2x_i \Sigma y - \Sigma x y - y_i \Sigma x) - \frac{B}{D} (2n x_i - 2 \Sigma x) \\ &= \frac{2x_i \Sigma y}{D} - \frac{\Sigma x y}{D} - \frac{y_i \Sigma x}{D} - \frac{2n B x_i}{D} + \frac{2B \Sigma x}{D} \end{aligned}$$

$$\frac{\partial B}{\partial x_i} = \left[\frac{2 \Sigma y - 2n B}{D} \right] x_i + \left[-\frac{\Sigma x}{D} \right] y_i + \left[\frac{2B \Sigma x - \Sigma x y}{D} \right] \quad (14)$$

La dérivée partielle de B par rapport à y_i est :

$$\frac{\partial B}{\partial y_i} = \frac{1}{D} \frac{\partial \beta}{\partial y_i} - \frac{\beta}{D^2} \frac{\partial D}{\partial y_i}$$

En utilisant les équations (13) et (9), on obtient :

$$\begin{aligned} \frac{\partial B}{\partial y_i} &= \frac{1}{D} (\Sigma x^2 - x_i \Sigma x) - \frac{\beta}{D^2} (0) \\ &= \frac{\Sigma x^2}{D} - \frac{x_i \Sigma x}{D} \end{aligned}$$

$$\frac{\partial B}{\partial y_i} = \left[-\frac{\Sigma x}{D} \right] x_i + \left[\frac{\Sigma x^2}{D} \right] \quad (15)$$

En utilisant les équations (14) et (15), nous pouvons désormais calculer les dérivées partielles de B par rapport à n'importe quel paramètre x_i et y_i . Ainsi, pour un graphique contenant n points (x, y) , l'erreur absolue δB associée à l'ordonnée à l'origine de la droite de régression linéaire peut être calculée en appliquant la **méthode différentielle** pour la propagation des erreurs :

$$\delta B = \left| \frac{\partial B}{\partial x_1} \right| \delta x_1 + \left| \frac{\partial B}{\partial x_2} \right| \delta x_2 + \dots + \left| \frac{\partial B}{\partial x_n} \right| \delta x_n + \left| \frac{\partial B}{\partial y_1} \right| \delta y_1 + \left| \frac{\partial B}{\partial y_2} \right| \delta y_2 + \dots + \left| \frac{\partial B}{\partial y_n} \right| \delta y_n$$