Annexe

Régression linéaire

Une série d'observations de n points (x_i,y_i) peut être approximée par une droite au sens des moindres carrés.
Soit $y = ax + b$ l'équation d'une droite. A chaque x_i, on peut calculer une valeur de y qui présente par rapport à la valeur observée y_i, un excès : $(y - y_i)^2$ au sens des écarts quadratiques.

L'excès total sur tous les points est

$$E = \Sigma (y_i - y)^2 = \Sigma (y_i - (a x_i + b))^2$$

L'ajustement consistera à trouver la droite de coefficient a et b qui minimise cet excès. Les variations de l'excès par rapport aux coefficients a et b, sont données par les dérivées partielles de E par rapport à a et b

$$\frac{\partial E}{\partial a} \quad et \quad \frac{\partial E}{\partial b}$$

L'excès E sera minimal lorsque les dérivées partielles s'annuleront.

$$\frac{\partial E}{\partial a} = 0 = \Sigma [-2 x_i (y_i - (a x_i + b))]$$

$$\frac{\partial E}{\partial b} = 0 = \Sigma [-2 (y_i - (a x_i + b))]$$

Il reste à résoudre un système à deux inconnues et deux équations :

$$\Sigma [x_i \,.\, y_i] - a \Sigma [x_i^2] - b \Sigma [x_i] = 0 \quad \Sigma [y_i] - a \Sigma [x_i] - n b = 0$$

En procédant par élimination, on obtient les coefficients a et b

$$a = \frac{n \Sigma x_i y_i - \Sigma x_i \Sigma y_i}{n \Sigma x_i^2 - (\Sigma x_i)^2} \qquad b = \frac{\Sigma y_i - a \Sigma x_i}{n}$$

On peut calculer le coefficient de corrélation linéaire

$$R = \frac{n \Sigma x_i y_i - \Sigma x_i \Sigma y_i}{\sqrt{(n \Sigma x_i^2 - \Sigma x_i \Sigma y_i)(n \Sigma y_i^2 - \Sigma x_i \Sigma y_i)}}$$

toujours positif, qui vaut 1 si la droite passe par tous les points, et moins que 1 autrement.

Pour obtenir les sigmas sur les coefficients a et b, on calcule la valeur intermédiaire de la variance

$$var = \frac{\Sigma y^2 + n b^2 + a^2 \Sigma x_i^2 - 2 (a \Sigma x_i y_i + b \Sigma y_i - a b \Sigma x_i)}{(n - 2)}$$

$$\sigma_a = \sqrt{\frac{n \,.\, var}{n \Sigma x_i^2 - \Sigma x_i^2}} \qquad \sigma_b = \sqrt{\frac{\Sigma x_i^2 \,.\, var}{n \Sigma x_i^2 - \Sigma x_i^2}}$$

Bevington P. R. - Data Reduction and Error analysis for the Physical Sciences - McGraw-Hill Book Company – 1969