

# TP 7 : Écart type et régression linéaire. Applications à la cinétique en Chimie.

## Exercice 1. Moyenne ; variance et écart-type d'une liste de nombres

On considère une liste L de nombres ; dans les formules qui suivront on notera  $n=\text{len}(L)$

**Question 1.** La moyenne d'une liste L de nombres est définie par :

$$m = \frac{1}{n} \sum_{i=0}^{i=n-1} L[i]$$

Écrire une fonction `moyenne(L)` qui prend en argument une liste L et qui retourne sa moyenne.

**Question 2.** Générer la liste `T1=[1,2,4,8,16,32,64,128,256,512,1024]` et tester votre fonction.

**Question 3.** La variance d'une liste L de nombres est définie par :

$$V(L) = \frac{1}{n} \sum_{i=0}^{n-1} (L[i] - m)^2$$

Écrire une fonction `variance(L)` qui prend en argument une liste L et qui retourne sa variance.

**Question 4.** Effectuer un test avec la liste T1

**Question 5.** L'écart-type d'une liste L de nombres est définie par :

$$\sigma(L) = \sqrt{V(L)}$$

Écrire une fonction `ecarttype(L)` qui prend en argument une liste L et qui retourne son écart-type.

**Question 6.** Effectuer un test avec la liste T1

## Exercice 2. Relation de König et sensibilité numérique

La littérature spécialisée cite souvent la relation de König, qui facilite les calculs de variance à la main. La relation de König donne la variance d'une liste L de moyenne m :

$$VP(L) = \frac{1}{n} \sum_{i=0}^{i=n-1} (L[i]^2 - m^2)$$

**Question 7.** Écrire une fonction `variancekonig(L)` qui prend en argument une liste L et qui retourne sa variance telle que calculée avec la relation de König.

**Question 8.** Effectuer un test de cette fonction avec la liste T1

**Question 9.** Définir les listes  $T2=[1, 2, 4, 8]$  et  $T3=[2^{27}+1, 2^{27}+2, 2^{27}+4, 2^{27}+8]$

Tester les calculs de variance avec ces listes avec les deux méthodes de calcul précédentes.  
Que constate-t-on ? Expliquer le phénomène.

### Exercice 3. Régression linéaire par la méthode des moindres carrés

La régression linéaire consiste à trouver pour une liste  $X=[x_0, \dots, x_{n-1}]$  et une liste  $Y=[y_0, \dots, y_{n-1}]$  de valeurs réelles, la « meilleure » droite  $y = ax + b$  approchant le nuage des points  $(x_i, y_i)_{0 \leq i < n}$ . On convient de choisir cette meilleure droite de manière optimale au sens des moindres carrés en ceci qu'elle minimise la somme des écarts au carré entre les  $y_i$  et les  $ax_i + b$ .

On note  $m_X$  et  $m_Y$  les moyennes respectives des listes  $X$  et  $Y$ .

La covariance de  $X$  et de  $Y$  est définie par 
$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=0}^{n-1} (x_i - m_X)(y_i - m_Y).$$

La variance de  $X$  est donc donnée par  $V(X) = \text{Cov}(X, X)$ .

Les coefficients  $a$  et  $b$  sont donnés par les formules suivantes :

$$a = \frac{\text{Cov}(X, Y)}{V(X)} \text{ et } b = m_Y - am_X$$

Le coefficient de corrélation linéaire, compris entre -1 et 1, mesure si la régression linéaire est de bonne qualité ou non. Un coefficient proche de 1 en valeur absolue dénote un bon ajustement ; plus ce coefficient est faible en valeur absolue et moins la régression linéaire est adaptée. Il est donné par la formule :

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

**Question 10.** Écrire une fonction `covariance(L1,L2)` qui prend en argument deux listes  $L1$  et  $L2$  et qui retourne leur covariance.

**Question 11.** Vérifier votre fonction en recalculant les variances de  $T1$  et  $T2$  à l'aide de la formule  $V(X)=\text{Cov}(X,X)$

**Question 12.** Écrire une fonction `reglin(L1,L2)` qui prend en argument deux listes  $L1$  et  $L2$  et qui renvoie les coefficients  $a, b$  et  $\rho$  obtenus par les formules de la régression linéaire par la méthode des moindres carrés.

**Question 13.** Effectuer un test avec  $L1=[0, 1]$  et  $L2=[0, 2]$  puis avec  $L1=[0, 0, 1, 1]$  et  $L2=[0, 1, 0, 1]$

**Exercice 4. Application à une cinétique chimique d'ordre 0**

On étudie dans les deux exercices suivants l'évolution de la concentration d'un réactif en fonction du temps pour une réaction chimique de la forme  $A+B \rightarrow C$ .

Dans certains cas, on peut écrire  $v = -\frac{d[A]}{dt}$  sous la forme  $k[A]^\alpha$

On dit alors que la réaction admet un ordre  $\alpha$  par rapport au réactif A.

Un exemple classique de réaction pharmaco-chimique d'ordre 0 est l'élimination de l'alcool dans le sang. On note  $f(t)$  le taux d'alcool dans le sang d'un sujet après son réveil du Nouvel An, où il n'a pas eu une consommation responsable, mais s'est arrêté de boire bien avant minuit, qui est associé à  $t = 0$ .

On fournit le relevé suivant :

| t (en h) | f(t) |
|----------|------|
| 0        | 2.3  |
| 1        | 2.17 |
| 3        | 1.84 |
| 5        | 1.56 |
| 7        | 1.27 |
| 10       | 0.81 |

**Question 14.** Vérifier que la réaction est bien d'ordre 0, en vérifiant que  $\rho \approx 1$

**Question 15.** Sur une même figure, représenter les points mesurés (on utilisera la fonction `plt.scatter`) et la droite obtenue avec la méthode des moindres carrés.

**Exercice 5. Exemple d'une réaction chimique d'ordre 1**

Certaines réactions sont d'ordre 1, on peut alors à l'aide de la méthode des moindres carrés déterminer une courbe approchée interpolant la concentration en fonction du temps. Cette interpolation ne sera pas toutefois optimale au sens de la méthode des moindres carrés, d'autres méthodes comme celle de Gauss-Newton donnant de meilleures interpolations.

La réaction de Landolt s'écrit :



Dans le cas d'un large excès d'ions I- et d'ions H+ la réaction de Landolt est d'ordre 1.

On étudie l'évolution de  $[\text{H}_2\text{O}_2]$  en fonction du temps. On pose  $f(t) = [\text{H}_2\text{O}_2](t) 10^{-4}$ .

Une étude expérimentale a fourni les valeurs suivantes :

| t (en s) | f(t) |
|----------|------|
| 0        | 296  |
| 29       | 281  |
| 58       | 266  |
| 91       | 251  |
| 127      | 236  |
| 170      | 221  |
| 203      | 207  |

**Question 16.** Justifier que  $\ln(f(t)) = At + B$ .

Construire des listes correspondant à  $\ln(f(t))$  et à  $t$

**Question 17.** En déduire une fonction approchant  $f(t)$ .

**Question 18.** Sur une même figure, représenter les points mesurés (on utilisera la fonction `plt.scatter`) et la courbe représentative déduite par précédemment.

### Exercice 6. Exo étoile : défaut de circularité

Les machines à mesurer tridimensionnelle (MMT) sont des instruments utilisés en métrologie dimensionnelle.

Elles permettent d'obtenir les coordonnées des points mesurés (palpés) sur une pièce mécanique. Ces coordonnées permettent de vérifier la validité dimensionnelle de la pièce, de vérifier que les cotes sont respectées. L'obtention de différentes coordonnées des points d'une pièce n'est pas très utile en soi. Une MMT est donc connectée à un ordinateur disposant d'un logiciel d'interprétation des coordonnées. Ce type de logiciel calcule les formes géométriques idéales aux points palpés sur les surface de la pièce à mesurer, ce qui autorise le calcul rapide d'un défaut de tolérance géométrique (forme, localisation, parallélisme, coaxialité...). Leur précision de mesure peut atteindre quelques nanomètres pour un volume de travail respectable d'une dizaine de centimètre de côté. ( source Wikipedia ) On se propose dans cet exercice de déterminer le cercle qui passe au mieux par un ensemble de points palpés situés dans un même plan , de manière à déterminer le défaut de circularité constaté.

Le cercle passant au mieux aura comme paramètres :

- son centre C de coordonnées  $(x_C, y_C)$
- son rayon R

Les  $n$  points palpés  $M_i$  ont pour coordonnées  $(x_i, y_i)$  avec  $1 \leq i \leq n$

En posant :

$$\begin{aligned}
 a &= \sum_{i=1}^n x_i^2 - \frac{1}{n} \left( \sum_{i=1}^n x_i \right)^2 \\
 b &= \sum_{i=1}^n x_i y_i - \frac{1}{n} \left( \sum_{i=1}^n x_i \right) \left( \sum_{i=1}^n y_i \right) \\
 c &= \sum_{i=1}^n y_i^2 - \frac{1}{n} \left( \sum_{i=1}^n y_i \right)^2 \\
 d &= \frac{1}{2} \left( \sum_{i=1}^n x_i^3 \sum_{i=1}^n x_i y_i^2 - \frac{1}{n} \sum_{i=1}^n x_i \left( \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 \right) \right) \\
 e &= \frac{1}{2} \left( \sum_{i=1}^n y_i^3 \sum_{i=1}^n y_i x_i^2 - \frac{1}{n} \sum_{i=1}^n y_i \left( \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 \right) \right)
 \end{aligned}$$

on trouve :

$$x_C = \frac{cd - be}{ac - b^2}$$

$$Y_C = \frac{ae - bd}{ac - b^2}$$

$$R = \sqrt{\frac{1}{n} \left( \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 - 2 \left( x_C \sum_{i=1}^n x_i + y_C \sum_{i=1}^n y_i \right) \right) + x_C^2 + y_C^2}$$

La machine à mesurer tridimensionnelle a enregistré une série de points  $M_i$  dont voici le relevé des coordonnées dans le repère machine sous forme de listes :

X= [13.648, 20.437, 29.113, 22.295, 27.992, 10.386, 24.542, 10.838, 19.698, 10.568]

Y= [27.064, 9.809, 24.078, 29.439, 25.5, 19.173, 28.981, 24.646, 10.11, 24.613]

**Question 19.** Écrire une fonction C( X,Y) qui prend en arguments les deux listes de coordonnées et qui renvoie les trois paramètres du cercle passant au mieux par les points mesurés.

**Question 20.** Compléter le script précédent de manière à afficher sur une même figure la position des points mesurés ainsi que le cercle qui passe au mieux (défini par la méthode des moindres carrés).