Fiche méthode : calculs d'incertitudes

Table des matières

I. Variabilité et incertitude-type **2**
1) La variabilité en sciences expérimentales . . . 2
2) Mesure et probabilités . . . 2
a) Densités de probabilités . . . 2
b) Espérance et écart type . . . 4
3) Incertitude-type relative . . . 5

II. Évaluation de $E(X)$ et de $u(X)$ en travaux pratiques **6**
1) Cas d'une mesure directe . . . 6
2) Cas d'une grandeur calculée à partir de grandeurs mesurées directement . . . 7
a) Utilisation des lois de la statistique . . . 7
b) Méthode de Monte-Carlo . . . 7
3) Formules de propagation des incertitudes . . . 8
a) Formule générale . . . 9
b) Cas d'une combinaison linéaire . . . 9
c) Cas d'un produit de puissances . . . 9
4) Le z-score . . . 10

III. Évaluation des incertitudes dans le cas d'un tracé de courbe **11**
1) Validation d'un modèle : polyfit . . . 11
2) Calcul de l'espérance et de l'incertitude par une méthode de Monte Carlo . . . 13
3) Utilisation d'une régression linéaire : coefficient de corrélation linéaire . . . 14

I. Variabilité et incertitude-type

1) La variabilité en sciences expérimentales

La mesure d'une grandeur est un processus généralement complexe qui entremêle de très nombreux processus. Cette complexité se traduit systématiquement par une variabilité de la mesure, qui implique que la répétition de la mesure conduit généralement à une valeur mesurée sensiblement différente de la première. Cette variabilité est naturelle et fait intrinsèquement partie de la mesure.

Cette variabilité peut provenir de nombreux aspects, dont les principaux sont les suivants :

- le choix de la méthode de mesure ;

  Exemple : Choisir de mesurer un petit élément à la règle graduée ou au pied à coulisse n'implique pas la même précision !

- les variations de l'environnement ;

  Exemple : Si l'on souhaite mesurer la célérité du son avec un protocole se déroulant sur une journée complète, comme la température de l'air va évoluer au cours du temps, la célérité du son aussi.

- les instruments de mesure ;

  Exemple : Mesurer une tension avec deux voltmètres semblant identiques amène parfois à une mesure de tension légèrement différente.

- le processus physique lui-même ;

  Exemple : Une expérience de mécanique quantique est intrinsèquement variable car la mécanique quantique ne prédit que des lois de probabilité.

- et surtout, la personne réalisant l'expérience.

  Généralement la personne réalisant l'expérience est la principale cause de variabilité de la mesure. Par ses gestes, ses choix et sa technique, cette personne introduit une variabilité importante. Il est donc totalement naturel que deux personnes réalisant la même expérience, dans les mêmes conditions, avec le même matériel, trouvent des valeurs différentes.

2) Mesure et probabilités

a) Densités de probabilités

Cette variabilité intrinsèque à la mesure entraîne qu'on fait appel à la théorie des probabilités pour décrire correctement la mesure d'une grandeur. On considère donc que la grandeur X à mesurer est une variable aléatoire dont les différentes réalisations possibles sont affectées de probabilités.

Dans le cas des grandeurs physiques et chimiques les différentes réalisations possibles peuvent prendre à priori toute valeur dans $\mathbb{R}$ et on est obligé de recourir à des *densités de probabilité* : on parle de variables **aléatoires continues** (contrairement aux variables aléatoires discrètes que vous avez vues en maths)

Définition

Il y a différents types de densités de probabilités. Parmi les lois modèles très souvent utilisées en physique-chimie on a :

- La **loi uniforme** sur un intervalle $[a, b]$.
- La **loi Gaussienne** (encore appelée loi normale) centrée sur une valeur x_m et d'écart type σ :

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\,\exp\left(-\frac{(x-x_m)^2}{2\sigma^2}\right)$$

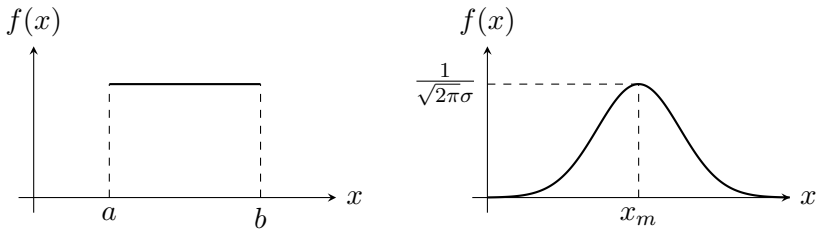


FIGURE 1 – Représentation des lois uniforme et gaussienne

- *Pour la loi uniforme il est certain que toute mesure d'une grandeur X donnera un résultat contenu dans l'intervalle $[a, b]$.*

La longueur du support $b - a$ est appelée **étendue**. On pose :

On aura donc la représentation de la Figure 2 :

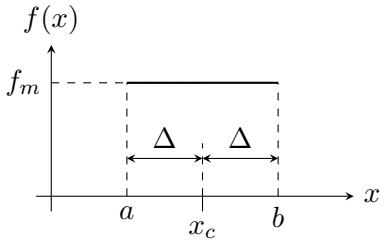


FIGURE 2 – Centre du support et demi étendue pour la loi uniforme

Exercice : déterminer f_m (cf. figure 2) en fonction de a et b.

- Au contraire la loi gaussienne $f(x)$ ne s'annulle jamais : si une grandeur X suit cette loi, il y a donc toujours une probabilité non nulle d'obtenir un résultat de mesure contenu dans n'importe quel intervalle $[c, d]$ de $\mathbb{R}$ même si c et d sont très éloignés de x_c.

b) Espérance et écart type

Dans le cas d'une variable aléatoire discrète X ne pouvant avoir que les réalisations x_1, x_2, ..., x_n affectées des probabilités $\mathscr{P}_1$, $\mathscr{P}_2$, ..., $\mathscr{P}_n$ on définit (cf. cours de maths) **l'espérance** $E(X)$ et **l'écart type** $u(X)$ par :

$$E(X) = \sum_{i=1}^{n} \mathscr{P}_i \, x_i \quad \text{et} \quad u(X) = \sqrt{\sum_{i=1}^{n} \mathscr{P}_i \times (x_i - E(X))^2}$$

Ces formules se généralisent à une variable aléatoire X affectée d'une densité de probabilité $f(x)$ de la façon suivante :

Dans le vocabulaire de la métrologie (science des mesures), l'écart type $u(X)$ s'appelle **l'incertitude type**. Elle caractérise donc la dispersion des mesures (c'est à dire leur variabilité).

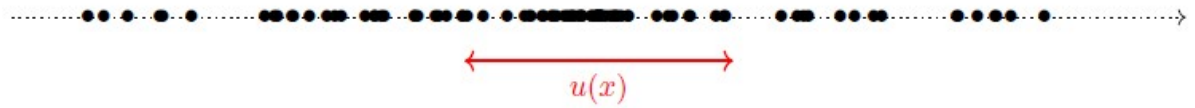


FIGURE 3 – Exemple de distribution de résultats de mesures x d'une grandeur X. On constate qu'en moyenne deux valeurs prises au hasard sont séparées au plus de quelques $u(X)$.

Résultats à connaître :

	Loi uniforme	**Loi Gaussienne**
E(X)	x_c	x_m
u(X)	$\frac{\Delta}{\sqrt{3}}$	σ

Idéalement, la mesure d'une grandeur X doit permettre de déterminer l'espérance $E(X)$ et l'incertitude type $u(X)$. Le résultat de mesure s'écrira sous la forme conventionnelle :

$$\boxed{X = E(X) \pm u(X)}$$

3) Incertitude-type relative

L'incertitude-type relative est définie par : $\dfrac{u(X)}{E(X)}$

Exercice :

1. Démontrer les valeurs de $E(X)$ et $u(X)$ pour une loi uniforme.
2. Même question pour la loi gaussienne (à faire chez soi). On donne les intégrales gaussiennes suivantes :

$$\int_{-\infty}^{+\infty} \exp(-\alpha u^2 + \beta u)\,\mathrm{d}u = \sqrt{\frac{\pi}{\alpha}}\,\exp\left(\frac{\beta^2}{4\alpha}\right)$$

et

$$\int_{-\infty}^{+\infty} u^2\,\exp(-\alpha u^2)\,\mathrm{d}u = \sqrt{\frac{\pi}{4\alpha^3}}$$

avec $\alpha \in \mathbb{R}_+^*$ et $\beta \in \mathbb{C}$.

II. Évaluation de $E(X)$ et de $u(X)$ en travaux pratiques

1) Cas d'une mesure directe

On suppose dans cette partie qu'on *mesure directement* la grandeur X : mesure d'une longueur avec une règle, d'un angle avec le vernier d'un goniomètre, d'une tension avec un voltmètre, etc...

Il y a deux façons de procéder :

a) **Cas d'une répétition de la mesure** :

Si on peut répéter la mesure plusieurs fois on peut utiliser les lois de la statistique. Si on réalise N mesures d'une grandeur X qui, du fait de la variabilité des résultats, ont donné les résultats x_1, x_2, ..., x_N alors :

$$\boxed{\overline{x} \overset{\text{déf}}{=} \frac{1}{N}\sum_{i=1}^{N} x_i \underset{N\gg 1}{\longrightarrow} E(X) \quad \text{et} \quad \sigma_{N-1} \overset{\text{déf}}{=} \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\overline{x})^2} \underset{N\gg 1}{\longrightarrow} u(X)} \qquad (1)$$

Cependant, pour que ce soit valable il faut que $N \gg 1$, ce qui est loin d'être le cas.

b) **Cas d'une mesure unique**

La plupart du temps il n'est pas possible de répéter la mesure et il faut se contenter d'un résultat unique.

Dans ce cas on suppose que la densité de probabilité $f(x)$ est à support borné et uniforme pour simplifier. On évalue alors expérimentalement **l'intervalle-support** $[x_c-\Delta, x_c+\Delta]$ dans lequel on estime qu'on a 100% de chances d'obtenir n'importe quel résultat de mesure si on répétait l'expérience un grand nombre de fois.

La demi-étendue Δ doit être prise la plus faible possible selon les critères personnels de l'expérimentateur et selon les conditions de l'expérience. *Il n'y a pas de règle générale.*

Exemple 1 :

On mesure une longueur L avec une règle graduée au millimètre : si on tombe directement sur une graduation, 14,7 cm par exemple, il est peut sembler naturel de prendre $L_c =$ 14,7 cm et $\Delta = 0{,}25$ mm.

Si la valeur est entre deux graduations, par exemple entre 14,7 cm et 14,8 cm, on prendra peut-être plus sûrement $L_c = 14{,}75$ cm et $\Delta = 0{,}5$ mm.

Enfin, une personne peu sûre d'elle peut choisir de prendre dans les mêmes cas $\Delta = 1$ mm.

Exemple 2 :

Pour les appareils de mesure, il est nécessaire de *consulter la notice de l'appareil.* Le constructeur indique souvent un pourcentage, par exemple 0,3% pour un voltmètre : on interprétera ce chiffre comme la demi-étendue exprimée en pourcents de la valeur lue.

On mesure par exemple une tension $u = 4{,}34$ V avec une précision constructeur de 0,3 %. On aura donc $\Delta = 4{,}3 \times 0{,}3/100 = 0{,}013$ qu'on s'empresse d'arrondir par valeur supérieure à un seul chiffre significatif : $\Delta = 0{,}02$ V. On aura donc $u_c = 4{,}34$ V et $\Delta = 0{,}02$ V

Exemple 3 :

Les composants électriques (résistances, condensateurs, inductances, ...) sont donnés avec une précision en pourcent (information constructeur). Par exemple sur une résistance, on peut lire R = 200 Ω ± 5 %. On aura donc $R_c = 200\ \Omega$ et $\Delta = 200 \times 5/100 = 10\ \Omega$.

Comme on a supposé la loi uniforme on a alors :

$$\boxed{E(X) = x_c \quad \text{et} \quad u(X) = \frac{\Delta}{\sqrt{3}}} \tag{2}$$

2) Cas d'une grandeur calculée à partir de grandeurs mesurées directement

a) Utilisation des lois de la statistique

Supposons que nous souhaitions évaluer une grandeur Z qui est une fonction de grandeurs mesurées directement X et Y : $Z = g(X, Y)$

Exemple : on a mesuré la température T et le volume V occupé par 1 mole d'un gaz parfait et on souhaite évaluer sa pression P au moyen de la formule $P = RT/V$.

X et Y étant des variables aléatoires, il en va de même de Z qui est donc caractérisée par un ensemble de réalisations possibles, affectées de probabilités. On souhaite donc évaluer $E(Z)$ et $u(Z)$.

Une fois encore, on a recours aux lois statistiques qui indiquent que si on mesure N fois X et Y avec des résultats :

$$(x_1, y_1), \ \ (x_2, y_2), \ \ ..., \ \ (x_N, y_N)$$

qui conduisent à N valeurs de Z :

$$z_1 = g(x_1, y_1), \ \ z_2 = g(x_2, y_2), \ \ ..., \ \ z_N = g(x_N, y_N)$$

alors dans la limite $N \gg 1$:

$$\boxed{\overline{z} = \frac{1}{N}\sum_{i=1}^{N} z_i = \frac{1}{N}\sum_{i=1}^{N} g(x_i, y_i) \underset{N\gg 1}{\longrightarrow} E(Z)} \tag{3}$$

et

$$\boxed{\sigma_{N-1}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - E(Z))^2 = \frac{1}{N-1}\sum_{i=1}^{N}(g(x_i, y_i) - E(Z))^2 \underset{N\gg 1}{\longrightarrow} u(Z)^2} \tag{4}$$

b) Méthode de Monte-Carlo

En pratique on utilise un logiciel comme Python faire le travail en tirant au sort $N \gg 1$ valeurs x_i et y_i dans leurs intervalles-précision suivant la loi de probabilité choisie à l'avance et on lui demande de calculer $E(Z)$ et $u(Z)$. Cette simulation informatique de la répétition d'expériences s'appelle **méthode de Monte-Carlo**.

Fonctions python pour réaliser le tirage au sort

On utilise la bibliothèque `numpy` pour simuler un processus aléatoire. On suppose qu'on l'a importée avec `import numpy as np`

Pour réaliser le tirage au sort on a les possibilités suivantes :

1. Si on suppose que les mesures de X et Y peuvent être uniforméments réparties sur tout leurs intervalles-précision on utilise la *loi uniforme* simulée par la fonction :

   `np.random.uniform(a,b)`

2. Si l'incertitude-type sur une des grandeurs est fournie (ce qui est très rare !) et qu'on pense qu'on peut utiliser une *loi gaussienne*, la fonction permettant de simuler cela est :

   `np.random.normal(xm,sigma)`

Exemple de code :

On a prélevé V_a mL d'un acide placés dans un becher et on en a réalisé le titrage par une base de concentration C_b que l'on a mis dans la burette. Le volume équivalent V_{eq} a été mesuré. On doit donc évaluer :

$$C_a = C_b \frac{V_{\text{eq}}}{V_a}$$

Les intervalles-précision ont été déterminés : V_a : [9,9 ; 10,1] mL, V_{eq} : [9,6 ; 9,8] mL et C_b : [0,099 ; 0,101] mol/L.

En ce qui concerne V_{eq} on pense que la valeur centrale est plus probable que les valeurs aux extrémités mais que ce n'est pas le cas pour C_b et V_a.

```
import numpy as np
N = 10000           # nombre d'expériences simulées
L = []              # liste pour contenir les résultats

for i in range(N) :
   Cb = np.random.uniform(0.099,0.101)
   Va = np.random.uniform(9.9,10.1)
   Veq = np.random.uniform(9.6,9.7,9.8)
   Ca = Cb*Veq/Va
   L.append(Ca)

ECa = np.mean(L)            # calcul de la valeur moyenne statistique (espérance)
uCa = np.std(L,ddof = 1)    # incertitude-type

print("concentration Ca :", ECa)
print("incertitude-type :", uCa)
```

concentration Ca : 0.09697633264752392
incertitude-type : 0.0008878909767078942

Conclusion :

On écrit le résultat en ne retenant qu'un seul (deux au maximum !) chiffres significatifs pour l'incertitude-type. On aura donc :

$$\boxed{C_a = (\ 0{,}0970 \pm 0{,}0009\ )\ \text{mol/L}}$$

3) Formules de propagation des incertitudes

Il existe aussi une autre façon d'obtenir $E(Z)$ et $u(Z)$ sans utiliser une méthode de Monte-Carlo. Cette technique, appelée technique de propagation des incertitudes, est parfois utilisée dans les sujets d'écrit de concours et il est donc nécessaire de la connaître. Les lois qui suivent sont tirées de théorèmes sur les variables aléatoires indépendantes dont nous ne parlerons pas ici.

a) Formule générale

Si $Z = g(X, Y)$ alors on a :

$$\boxed{E(Z) = g\,[\,E(X), E(Y)\,]}$$

et

$$\boxed{u(Z) = \sqrt{\left(\frac{\partial g}{\partial X}\right)^2 (u(X))^2 + \left(\frac{\partial g}{\partial Y}\right)^2 (u(Y))^2}}$$

où les deux dérivées partielles doivent être prises en $E(X)$ et $E(Y)$.

Il y a deux cas particuliers très importants en pratique qu'il faut connaître. Vous pouvez les utiliser directement sans avoir à les "redémontrer" à chaque fois.

b) Cas d'une combinaison linéaire

Dans le cas où la grandeur à évaluer Z s'exprime sous la forme d'une combinaison linéaire c'est à dire :

$$\boxed{Z = \alpha\, X + \beta\, Y} \tag{5}$$

où α et β sont des coefficients numériques parfaitement connus (par exemple des entiers ou bien π dont on peut donner autant de chiffres significatifs qu'on veut), avec comme **cas particuliers** la somme $Z = X + Y$ ou la différence $Z = X - Y$. On a [1] :

$$\boxed{E(Z) = \alpha\, E(X) + \beta\, E(Y) \quad \text{et} \quad u(Z) = \sqrt{\alpha^2\,(u(X))^2 + \beta^2\,(u(Y))^2}} \tag{6}$$

Démo :

Exemple : on souhaite mesurer une distance $D = d_1 - d_2$. On a estimé par mesure directe : $d_1 = 10{,}3 \pm 0{,}2$ mm et $d_2 = 2{,}7 \pm 0{,}1$ mm.

c) Cas d'un produit de puissances

Dans le cas où la grandeur à évaluer Z s'exprime sous la forme suivante :

$$\boxed{Z = k\, X^{\alpha} \times Y^{\beta}} \tag{7}$$

1. La généralisation de ce résultat à une combinaison linéaire de plus de deux termes est immédiate.

où k, α et β sont des coefficients numériques parfaitement connus, avec comme **cas particulier** un simple produit $Z = X \times Y$ ou un quotient $Z = X/Y$. La loi s'écrit :

$$\boxed{E(Z) = k\,E(X)^{\alpha} \times E(Y)^{\beta} \quad \text{et} \quad \frac{u(Z)}{E(Z)} = \sqrt{\alpha^2 \left(\frac{u(X)}{E(X)}\right)^2 + \beta^2 \left(\frac{u(Y)}{E(Y)}\right)^2}} \tag{8}$$

Démo :

Exemple : on souhaite mesurer la vitesse des ultrasons dans l'air à l'aide de la relation $c = \lambda f$. On a estimé par mesure directe : $\lambda = (8{,}3 \pm 0{,}6)$ mm et $f = (40.10^3 \pm 58)$ Hz.

Remarque :

Les physiciens n'écrivent jamais $E(c) = 332$ m.s^{-1} ou $E(\lambda) = 8{,}3$ mm. Au lieu de cela ils écrivent de façon plus naturelle (mais moins précise quant à l'objet mathématique manipulé) $c = 332$ m.s^{-1} ou bien $\lambda = 8{,}3$ mm

4) Le z-score

Supposons qu'on ait mesuré une grandeur X, c'est à dire qu'on connaît $E(X)$ et $u(X)$. De plus on sait par des informations dans des tables ou sur internet, quelle est la valeur attendue x_a et son incertitude-type $u(x_a)$.

Afin de pouvoir comparer les deux résultats, on calcule le **z-score** :

$$\boxed{\text{z-score} = \frac{|E(X) - x_a|}{\sqrt{u(X)^2 + u(x_a)^2}}} \tag{9}$$

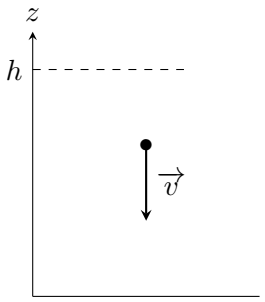
On considère que le résultat de mesure $E(X)$ est **compatible** avec la valeur attendue si z-score $\leqslant 2$. Ce seuil à 2 est d'origine historique. On le retrouve dans de nombreux champs scientifiques, comme la médecine, la pharmacie, la biologie, la psychologie, l'économie, etc...

III. Évaluation des incertitudes dans le cas d'un tracé de courbe

1) Validation d'un modèle : polyfit

Exemple 1 :

Supposons que nous souhaitions mesurer l'accélération de la pesanteur g à l'aide d'une expérience de chute libre. On laisse tomber un objet d'une hauteur h et on enregistre une vidéo de la chute.



En supposant une vitesse initiale v_0 non nulle (même si on s'arrange pour qu'elle soit très petite) l'équation horaire du mouvement s'écrit :

$$z(t) = -\frac{gt^2}{2} + v_0 t + h \tag{10}$$

Un pointage de la vidéo a donné les résultats suivants :

z (cm)	200	197	190	177	160	137	110	77,0
t (ms)	0	72	143	214	286	357	429	500

Comment valider le modèle théorique (10) à l'aide de cette série de mesures ?

Utilisation de la fonction `polyfit` de la bibliothèque `numpy`

Pour commencer, la fonction `polyfit` permet de récupérer les valeurs des coefficients d'un polynôme de degré n quelconque $z = p_0 t^n + p_1 t^{n-1} + ... + p_{n-1} t + p_n$ qui modélise au mieux deux séries de données expérimentales sous la forme de listes (ou de vecteurs numpy) `T` et `Z`. Sa signature est (on suppose qu'on a écrit `import numpy as np`) :

`np.polyfit(T:np.array, Z:np.array, n:int)` $\longrightarrow$ `np.array`

où `n` est le degré du polynôme. La valeur renvoyée est un tableau `p` de longueur n dont les valeurs `p[i]` sont les coefficients du polynôme qui modélise au mieux la série de données [2].

2. Attention : p[0] est le coefficient de t^n, monôme de degré le plus élevé du polynôme.

Exemple de code :

```
import numpy as np
import matplotlib.pyplot as plt

T = np.array([0,72e-3,143e-3,214e-3,286e-3,357e-3,429e-3,500e-3])
Z = np.array([2.00,1.97,1.90,1.77,1.60,1.37,1.10,0.77])

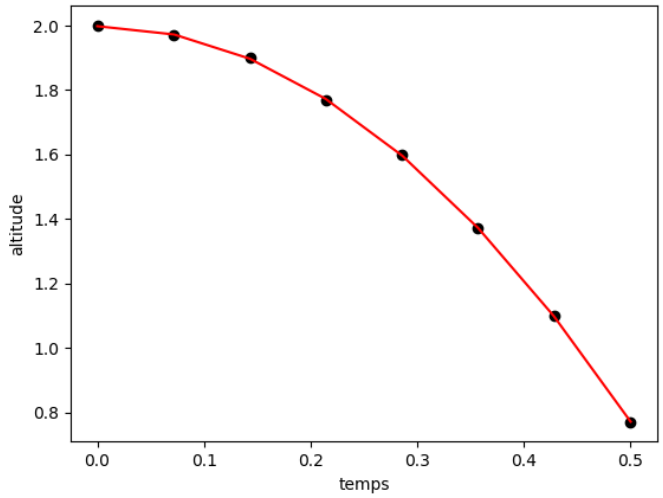
p = np.polyfit(T,Z,2)     # modélisation par polynôme degré 2
print(p[0])
print(p[1])
print(p[2])
print("g = ",-2*p[0])

Zmodele = p[0]*(T**2) + p[1]*T + p[2]      # polynôme modèle de degré 2

# Représentation graphique des résultats

plt.figure()
plt.plot(T,Z,"ko")     # résultats expérimentaux : points noirs
plt.plot(T,Zmodele, "red")      # polynôme modèle en rouge
plt.xlabel("temps")
plt.ylabel("altitude")
plt.show()
```

-4.902383533467539
-0.004596653619921438
1.9985998764592394
g = 9.804767066935078



Exemple 2 :

Le monoxyde d'azote est oxydé selon l'équation-bilan : 2 NO + O_2 = 2 NO_2

Il s'agit d'une réaction totale dans les conditions de l'expérience et on a relevé les résultats suivants en suivant une méthode de dégénérescence de l'ordre : [NO] $\gg$ [O_2] :

t (en s)	10	20	30	60	120	340	360
[O_2] (en mol.L^{-1})	$9{,}3.10^{-6}$	$8{,}6.10^{-6}$	$8{,}0.10^{-6}$	$6{,}4.10^{-6}$	$4{,}1.10^{-6}$	$1{,}7.10^{-6}$	$0{,}70.10^{-6}$

On veut montrer que l'ordre partiel par rapport à O_2 vaut 1. Il faut donc montrer que :

$$\ln([O_2] = -k_{app}\, t + \ln([O_2]_0)$$

```
import numpy as np
import matplotlib.pyplot as plt

T = np.array([10,20,30,50,60,120,240,360])
C = np.array([9.3e-6,8.6e-6,8.0e-6,6.4e-6,4.1e-6,1.7e-6,0.70e-6])
lnC = np.log(C)     # logarithme des concentrations

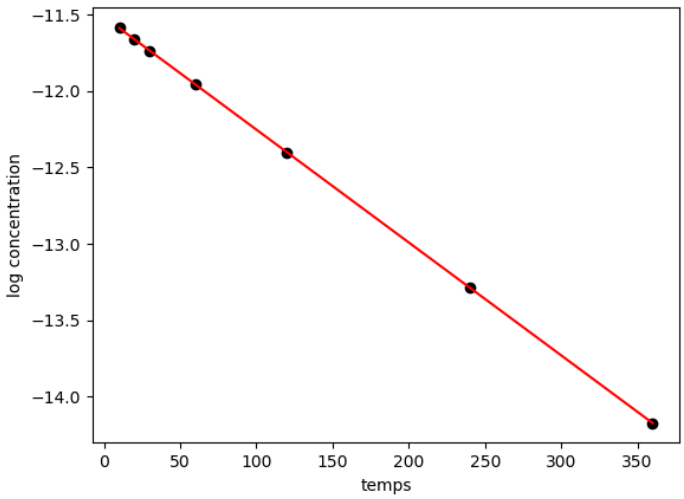
p = np.polyfit(T,lnC,1)      # polynôme modèle de degré 1
print(p[0])
print(p[1])
print("kapp = ",-p[0])

lnCmodele = p[0]*T + p[1]      # droite modèle

# Représentation graphique des résultats

plt.figure()
plt.plot(T,lnC,"ko")     # résultats expérimentaux : points noirs
plt.plot(T,lnCsim, "red")      # droite modèle en rouge
plt.xlabel("temps")
plt.ylabel("log concentration")
plt.show()
```

-0.007381210008831938
-11.515128726004752
kapp = 0.007381210008831938



2) Calcul de l'espérance et de l'incertitude par une méthode de Monte Carlo

Reprenons l'exemple 1. La représentation graphique nous indique que le modèle est bon, mais cela ne nous renseigne pas sur les incertitudes. Supposons que la demi-étendue sur le temps soit $\Delta_t = 1$ ms et que celle sur z soit $\Delta_z = 2$ mm. Quelle est alors l'incertitude-type $u(g)$ sur l'accélération de la pesanteur g et quelle est l'espérance $E(g)$?

Pour cela il faut élargir l'utilisation de la fonction `np.random.uniform` (ou de `np.random.triangular`). Un paramètre supplémentaire de la fonction est `size` Par défaut la fonction ne renvoie qu'une seule valeur. Si on met `size = n` alors la fonction renvoie un vecteur numpy ayant n composantes, toutes tirées au sort selon la loi uniforme (ou triangulaire).

La démarche est donc la suivante :

- L'expérience nous a donné K couples de valeurs (t_k, z_k) avec leurs demi-étendues Δ_t et Δ_z.
- On tire alors au hasard K valeurs appartenant aux domaines :

$$[t_k - \Delta_t, t_k + \Delta_t] \times [z_k - \Delta_z, z_k + \Delta_z]$$

- On utilise la fonction `np.polyfit` avec ces K valeurs, ce qui nous renvoie un vecteur `p`.
- On stocke `-2*p[0]` (ce qui est censé donner g) dans une liste `L` et on recommence N fois avec $N = 1000$ ou 10 000 (ou plus, cela dépend de la rapidité de votre ordinateur).
- On calcule `np.mean(L)` et `np.std(L,ddof = 1)`

Exemple de code :

```
import numpy as np

T = np.array([0,72e-3,143e-3,214e-3,286e-3,357e-3,429e-3,500e-3])
Z = np.array([2.00,1.97,1.90,1.77,1.60,1.37,1.10,0.77])

Delta_t = 1e-3      # demi-étendue sur le temps
Delta_z = 2e-3      # demi-étendue sur l'altitude

N = 10000      # nombre de tirages au sort
L = []
nbrePoints = len(T)      # nombre de points de mesure: len(T) == len(Z)

for i in range(N) :
   Tsim = T + np.random.uniform(-Delta_t,Delta_t,size = nbrePoints)
   Zsim = Z + np.random.uniform(-Delta_z,Delta_z,size = nbrePoints)
   p = np.polyfit(Tsim,Zsim,2)
   L.append(-2*p[0])      # stocker la valeur de g

Eg = np.mean(L)
ug = np.std(L, ddof = 1)

print("accélération pesanteur : ",Eg)
print("incertitude type : ",ug)
```

accélération pesanteur : 9.793070571658339
incertitude type : 0.11375554355244688

On a donc : $g = (\ 9{,}79 \pm 0{,}12\ )\ \text{m.s}^{-2}$

3) Utilisation d'une régression linéaire : coefficient de corrélation linéaire

Reprenons l'exemple 2. On a vu qu'on obtenait une droite en représentant $\ln([O_2])$ en fonction de t, ce qui valide l'ordre 1 pour le réactif O_2. Un argument supplémentaire en faveur de cela est de calculer le coefficient de corrélation $|r|$ (sa valeur absolue) de la droite obtenue (méthode de régression linéaire).

Avec numpy cela se fait avec la fonction `np.corrcoef` dont la signature est :

`np.corrcoef(X: np.array, Y : np.array)` $\longrightarrow$ `np.array`

Cette fonction prend en argument deux tableaux numpy (ou deux listes) `X` et `Y` qui représentent les valeurs expérimentales (x_i, y_i) dont on veut s'assurer qu'elles sont alignées sur une droite.

Elle renvoie une matrice carrée 2×2, appelée **matrice covariance** qui est symétrique et dont les *élément non-diagonaux* sont justement le coefficient de corrélation cherché.

Exemple de code :

```
import numpy as np

T = np.array([10,20,30,50,60,120,240,360])
C = np.array([9.3e-6,8.6e-6,8.0e-6,6.4e-6,4.1e-6,1.7e-6,0.70e-6])
lnC = np.log(C)     # logarithme des concentrations

M = np.corrcoef(T,lnC)      # matrice covariance

print(M)
print("coefficient de corrélation : ", M[0,1])
print("coefficient de corrélation : ", M[1,0])
```

```
[[ 1.         -0.99999732]
 [-0.99999732  1.        ]]
coefficient de corrélation : -0.9999973163644661
coefficient de corrélation : -0.9999973163644661
```

Bilan :

Le calcul du coefficient de corrélation *vient renforcer l'argumentation* d'une relation affine entre deux grandeurs X et Y. Il se fait **après** avoir tracé la courbe de Y en fonction de X, tracé qui doit montrer qu'une loi linéaire (affine) est un bon modèle. Il **ne peut pas se substituer** au tracé de cette courbe (certaines courbes qui ne sont manifestement pas des droites donnent aussi un très bon coefficient de corrélation).

Enfin, le calcul du coefficient de corrélation valide un modèle de relation affine entre X et Y), mais *cela n'a rien à voir avec les calculs d'incertitude.*