Mesures et Incertitudes

A completer : https://colab.research.google.com/drive/1VRXCPJy9ULHmzaRzusse-lfLhS4L5ZQm?usp=sharing

Correction : https://colab.research.google.com/drive/149_LpAwVUJCd9PUZtIn9tju_4Jm7mN3Y?usp=sharing

I. Introduction à la mesure et à sa variabilité

- **Mesure d'une grandeur physique**

En physique-chimie, on appelle mesure une procédure expérimentale qui conduit à attribuer un ensemble de valeurs numériques à une grandeur, accompagné d'une unité appropriée. Par exemple, la mesure de la longueur d'une feuille de papier A4 à l'aide d'une règle conduit à un ensemble de valeurs numériques, associées à l'unité cm ou mm

Le résultat d'une mesure décrit cet ensemble de valeurs en le complétant par des explications sur la manière dont elles ont été obtenues

- **Variabilité d'une mesure**

La mesure est intrinsèquement variable : bien qu'on ne s'en aperçoive pas toujours, si on répète une mesure, on trouve souvent une valeur numérique différente. Par souci de clarté, on peut appeler chacune de ces valeurs numériques observation.

Exemple : On mesure le volume équivalent lors d'un titrage acide-base. C'est une observation. Si on recommence, on va probablement obtenir une autre valeur. Dans une classe, on peut rassembler l'ensemble des observations : elles constituent une mesure.

II. Représentation de la variabilité d'une série de mesures : moyenne et écart-type

Deux grandeurs permettent de décrire un ensemble d'observations de manière le plus réduit possible :

1. la **moyenne** arithmétique des observations, qu'on appelle parfois « moyenne expérimentale »
2. l'**écart-type** des observations, qu'on qualifie parfois d'expérimental, ou d'échantillon.

En notant les N observations x_i,

- $\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$, correspond à la **valeur moyenne expérimentale** des N observations
- $s_x = \sqrt{\frac{1}{(N-1)}\sum_{i=1}^{N}(x_i - \overline{x})^2}$ correspond à l'**écart-type expérimental** des N observations

Python : **Valeur moyenne :** **xbar = numpy.average(…)**
Ecart-type expérimentale : **sx = numpy.std(…, ddof = 1)**

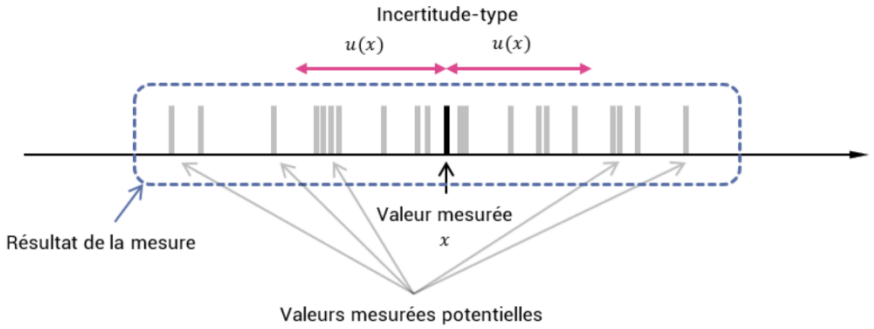
III. Valeur mesurée et incertitude-type

- La **valeur mesurée** est la meilleure estimation possible de la grandeur mesurée.

Lorsqu'on répète des observations et que l'on constate qu'elles varient, on choisit leur **moyenne** arithmétique comme meilleur estimateur.

Si on n'a qu'une valeur unique, alors c'est celle qu'on prend.

- L'**incertitude-type** est l'estimation à l'aide d'un écart-type, de la dispersion des valeurs raisonnablement attribuables à la grandeur mesurée. La **valeur mesurée** est une de ces valeurs. L'**incertitude-type** quantifie donc la variabilité *potentielle* de cette unique **valeur mesurée**. En cela elle reflète son **incertitude**.



IV. Comparaison de deux valeurs et écart normalisé

- **Comparaison entre deux valeurs x_1 et x_2 entachées d'incertitude $u(x_1)$ et $u(x_2)$**

  L'écart normalisé (aussi appelé z-score) permet de comparer ces deux valeurs. Il s'écrit alors :

$$z = \frac{|x_1 - x_2|}{\sqrt{u(x_1)^2 + u(x_2)^2}}$$

  si $z < 2$: les deux valeurs sont **compatibles** entre elles

  si $z \geq 2$: les deux valeurs sont **incompatibles** entre elles

- **Comparaison avec une valeur de référence dont on néglige l'incertitude**

  Pour valider un résultat expérimental, la valeur mesurée est parfois comparée à une valeur de référence. Si l'incertitude de la valeur de référence est négligée, l'écart normalisé devient :

$$z = \frac{|x - x_{ref}|}{u(x)}$$

  si $z < 2$: le résultat de la mesure est **compatible** avec la valeur de référence

  si $z \geq 2$: il y a **incompatibilité** entre la valeur de la mesure et la valeur de référence

V. Évaluer une incertitude-type par une approche statistique (type A)

Incertitude-type associée à la moyenne unique de N observations

$$u(\overline{x}) = \frac{s_x}{\sqrt{N}}$$

Exemple n°1 : Absorbance

Une série de N = 24 mesures de l'absorbance d'une solution de sulfate de cuivre donne :

essai	1	2	3	4	5	6	7	8	9	10	11	12
A	0,953	0,945	0,967	0,985	0,963	0,968	0,954	0,971	0,957	0,964	0,968	0,983

essai	13	14	15	16	17	18	19	20	21	22	23	24
A	0,958	0,941	0,963	0,988	0,966	0,962	0,952	0,974	0,954	0,971	0,969	0,981

1. Déterminer la valeur moyenne de $\overline{A}$
2. Déterminer l'écart-type expérimental s_A
3. Déterminer l'incertitude-type : $u(A) = \frac{s_A}{\sqrt{N}}$

Remarque : Les calculs de la valeur moyenne et de l'écart-type permettent d'éliminer les points « aberrants », que l'on prend situés à plus de deux écarts-types de la moyenne. Ici ce sont tous les points hors de l'intervalle, dit de confiance, compris entre 0,9649-2x0,01209136 = **0,940** et 0,9649+2x0,01209136 = **0,989**. Dans notre cas, tous les points sont acceptés.

4. Ecriture du résultat : donner la valeur de l'absorbance avec son incertitude
   On exprime généralement l'incertitude – type avec deux chiffres significatifs imposant l'écriture de la valeur mesurée :

VI. Évaluer une incertitude-type par une autre approche (type B)

On se place ici dans le cas où on n'a pu réaliser **qu'une seule observation**, **ou bien dans le cas où la répétition des observations conduit exactement à la même valeur**.

Exemple : On mesure la longueur d'une feuille A4 avec un triple-décimètre gradué au millimètre et on trouve 297 mm. Si on recommence, on trouve encore 297 mm

- **Observations équiréparties dans un intervalle** $[m - a\,;m + a]$: $\boldsymbol{u(x) = \frac{a}{\sqrt{3}}}$

- **Incertitude-type associée à un appareil de mesure :**
  L'évaluation des incertitudes sur une mesure directe est généralement effectuée à partir de données fournies par le constructeur de l'appareil de mesure.

Type d'incertitude donnée par le constructeur	Incertitude type	Exemples
$\Delta_c = \pm \ldots$	$u(x) = \frac{\Delta_c}{\sqrt{3}}$	**Résistance donnée à +/- 5% près** pour une résistance de 13,4Ω : $\frac{\Delta_c}{\sqrt{3}} = \cdots \ldots \ldots = \cdots \ldots \ldots$
$\Delta_c = \ldots$	$u(x) = \frac{\Delta_c}{2\sqrt{3}}$	**Balance précise au mg** soit Δ_c = 1 mg ; l'incertitude-type sur chaque pesée (dont la tare) est : $\frac{\Delta_c}{2\sqrt{3}} = \cdots \ldots \ldots$

Incertitude-type composée : à l'aide d'une formule mathématique

S'il existe plusieurs causes d'erreurs indépendantes, d'incertitude-type u_i, pour une grandeur x mesurée directement ou indirectement, l'incertitude type $u(x)$ de la mesure est dite **composée** et on l'estime par la **formule de propagation des incertitudes**.

- **Evaluation sur une mesure directe**

  La formule de propagation des incertitudes sur une mesure directe : $\boldsymbol{u(x) = \sqrt{\sum_i (u_i)^2}}$

- **Evaluation sur une mesure indirecte**

Lorsque la grandeur **X** n'est pas mesurée directement mais est calculée à partir de **plusieurs grandeurs A, B, C, …**, mesurées elles directement, la formule de propagation des incertitudes dépend de la forme mathématique de la relation existant entre **G** et **A**, **B**, **C**, … :

Cas d'une somme ou d'une différence **Exemple :** $\boldsymbol{X = A + B - C}$	$\boldsymbol{u(X) = \sqrt{u(A)^2 + u(B)^2 + u(C)^2}}$
Cas d'un produit ou d'un quotient **Exemple :** $\boldsymbol{X = \frac{A \times B}{C}}$	$\boldsymbol{u(X) = X\sqrt{\left(\frac{u(A)}{A}\right)^2 + \left(\frac{u(B)}{B}\right)^2 + \left(\frac{u(C)}{C}\right)^2}}$

Exemple n°2 : Concentration en quantité de matière acide éthanoïque (1)

Un volume V_A = 10,0mL, prélevé à l'aide d'une pipette de 10mL de classe A (+/- 0,02mL), de solution aqueuse d'acide éthanoïque de concentration C_A inconnue est dosé par de la soude de concentration $C_{soude} = 0{,}10\ mol.L^{-1}$ (à 10%). Le volume équivalent mesuré (par suivi colorimétrique) est de V_e = 10,3mL à l'aide d'une burette graduée de classe A (+/- 0,05mL) de pas 0,1 mL. L'équivalence est détectée à une goutte près (Vgoutte = 0,05 mL). En déduire la concentration C_A.

1. Déterminer la concentration C_A en acide éthanoïque
2. Exprimer l'incertitude $u(C_A)$ en fonction de $u(C_{soude})$, $u(V_e)$ et $u(V_A)$
3. Donner l'expression de $u(C_{soude})$ et $u(V_A)$
4. Exprimer l'incertitude $u(V_e)$ en fonction de $u_{burette}$, u_{goutte} et u_{pas}. Donner l'expression de $u_{burette}$, u_{goutte} et u_{pas}.
5. Ecriture du résultat : donner la concentration en acide éthanoïque avec son incertitude

En pratique :

Plutôt que d'engager de longs calculs, il est attendu une véritable réflexion sur les sources d'incertitude, et donc d'évaluer le plus simplement (et rapidement) possible l'ordre de grandeur de l'incertitude (en négligeant, si besoin certaines sources d'incertitude devant les autres).

Exemple 1 : lors d'une mesure d'un temps au chronomètre, l'incertitude dominante n'est pas le centième de seconde correspondant à la précision du chronomètre mais le temps de réaction de l'expérimentateur !

Exemple 2 : L'incertitude sur la température d'un bloc de métal sorti d'un congélateur pour être introduit dans un calorimètre ne correspond pas à au demi degré de précision donné par le thermomètre dans le congélateur....

Exemple 3 : l'incertitude sur la fréquence de résonance ne correspond pas à celle donnée sur la fréquence délivrée par le GBF mais à la méthode de pointage de la résonance !

Incertitude-type composée : à l'aide de la méthode Monte-Carlo

Une autre méthode qui permet d'évaluer l'incertitude-type (B), lorsque l'expérience n'est pas répétée plusieurs fois est une méthode numérique, appelée méthode de Monte-Carlo. Celle-ci consiste à simuler numériquement un grand nombre d'expériences. L'écart-type de l'ensemble des valeurs obtenues lors de ces répétitions fournit l'incertitude-type recherchée.

Un logiciel tel que python est nécessaire pour créer un grand nombre de valeurs tirées au sort dans l'intervalle représentant le résultat de la mesure, avec une distribution uniforme ou normale. Une telle simulation est avantageuse car elle permet d'éviter des calculs fastidieux de propagation d'incertitudes. Il faut la privilégier autant que possible lorsqu'il y a plusieurs sources d'incertitude que l'on ne peut négliger.

Python : générer des nombres aléatoires

Si l'on connaît l'incertitude-type (u(X)	Si l'on suppose qu'aucune des valeurs dans l'intervalle n'est plus probable qu'une autre (demi-étendue a)	Si la valeur centrale de l'intervalle est plus probable que les valeurs aux extrémités de l'intervalle (demi-étendue a)
distribution gaussienne	**distribution uniforme**	**distribution triangulaire**
numpy.random.normal(X,u(X)) **tirage aléatoire de N valeurs :** **numpy.random.normal(X,u(X),N)**	**numpy.random.uniform(X-a,X+a)** **tirage aléatoire de N valeurs :** **numpy.random.uniform(X-a,X+a,N)**	**numpy.random.triangular(X-a,X+a)** **tirage aléatoire de N valeurs :** **numpy.random.triangular(X-a,X+a,N)**

Les 4 étapes du code sont :

1. Définir le nombre de tirages à simuler.
2. Réaliser le tirage de valeurs "expérimentales".
3. Calculer, avec chaque jeu de valeurs simulées, la grandeur à déterminer et stocker cette valeur dans une liste.
4. Calculer la moyenne et l'écart-type des valeurs de la liste.

Exemple n°3 : Concentration en quantité de matière acide éthanoïque (2)

Sur l'exemple précédent, estimer l'incertitude sur la concentration en acide éthanoïque par la méthode Monte-Carlo. Donner alors la concentration en acide éthanoïque avec son incertitude

VII. Régression linéaire

Une régression linéaire permet de vérifier la compatibilité entre un modèle linéaire choisi et des données expérimentales. Un accord ne signifiera pas pour autant que le modèle choisi est valide.

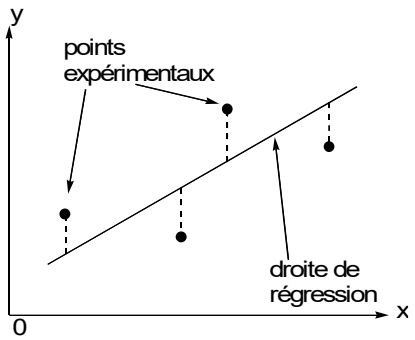
- **Principe**

→ **Couple de valeurs expérimentales** : (x_i, y_i)

→ **Modèle linéaire envisagé** : $y = ax + b$

→ **Obtention de *a* et *b***: en minimisant la quantité $\sum_{i=1}^{n}\left[y_i - (ax_i + b)\right]^2$

(méthode des « moindres carrés »)



- **Exploitation du coefficient de corrélation (R^2)**

Si $R^2 \rightarrow 1$, cela permet d'avoir une première idée de la compatibilité entre les données expérimentales et un modèle choisi. Mais, pour arriver à cette première conclusion, il faut toujours visualiser la droite de régression et les points expérimentaux pour repérer d'éventuelles erreurs de saisie ou que les points sont plutôt sur une courbe.

- **Visualisation graphique**

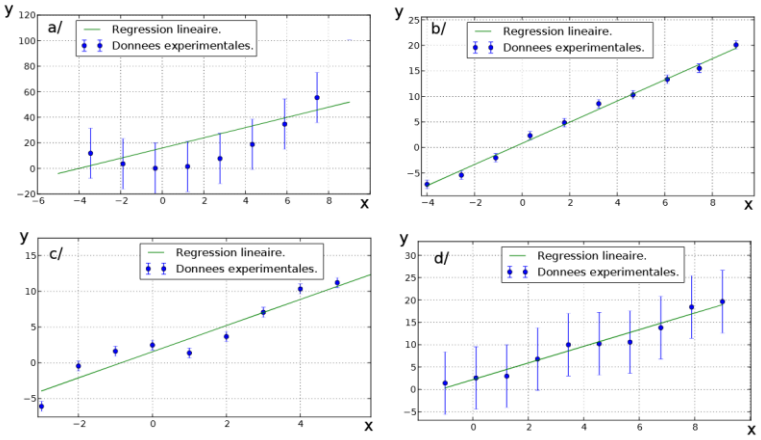
Python :	**Données en abscisse :**	**X=numpy.array([..])**
	Données en ordonnées :	**Y=numpy.array([..])**
	Régression linéaire :	**a,b = numpy.polyfit(X,Y,1)**
	Tracé de la droite de régression linéaire :	**regression = a*X+b**
	Coefficient directeur :	**a**
	Ordonnée à l'origine :	**b**
	Affichage des résidus (écart entre les points expérimentaux et la droite de régression) :	**residus=Y-(a*X+b)**

- **Affichage de la variabilité de chaque point**

Pour plus de rigueur, il convient d'évaluer la variabilité de chaque point à l'aide des incertitudes.

Python :	**Affichage des points avec barres d'erreur**	**plt.errorbar(X,Y,yerr=…,xerr=….)**

- **Les différents cas**



- Cas **a** : la droite de régression passe par les barres d'erreur, mais cependant on note une nette courbure : le modèle linéaire n'est donc certainement pas bon.
- Cas **b** : pas de tendance particulière, et la droite de régression passe par les barres d'erreur : on valide le modèle linéaire.
- Cas **c** : la droite de régression ne passe pas par les barres d'erreur : on rejette le modèle. (il faudra vérifier qu'on n'a pas pris des incertitudes trop petites)
- Cas **d** : pas de tendance nette et la droite passe par les barres d'erreur : on valide le modèle. Remarquons toutefois que l'incertitude est grande et que d'autres modèles, non linéaires, pourraient aussi convenir : la conclusion n'est donc pas très forte. Il faut aussi vérifier qu'on n'a pas surestimé les incertitudes.

- **Compatibilité entre la droite – modèle et les points expérimentaux**

> **Le modèle proposé sera compatible avec les données expérimentales si :**
>
> - **les points expérimentaux s'alignent bien suivant la droite-modèle**
> - **les résidus sont répartis aléatoirement autour de la droite-modèle**
> - **la valeur absolue des résidus est négligeable devant l'étendue des valeurs mesurées**

Exemple n°4 : Détermination de la concentration inconnue en glucose d'une solution par étalonnage (polarimétrie)

On souhaite déterminer la concentration inconnue en glucose dans une solution S par étalonnage. On mesure pour cela par polarimétrie l'angle de déviation α pour cinq solutions de glucose de concentration en masse C préalablement préparée :

α (en °)	2,15	4,0	5,5	6,5	8,3
C (en g/L)	10	20	30	40	50

On mesure l'angle de déviation de la solution S de glucose de concentration inconnue : $\alpha_{mesuré} = 3{,}1°$

Partie 1 : Tracé, droite de régression, résidus

1. Tracé α en fonction de C

2. Réaliser une régression linéaire à l'aide de la fonction *polyfit* de la bibliothèque *numpy (np)*
   a,b = np.polyfit(x,y,1) permet d'effectuer une régression linéaire sur les listes x et y.
   Elle retourne le meilleur couple a, b, tel que la droite d'équation y = ax + b soit au plus près des points (xi,yi)

3. Faire afficher la droite de régression ainsi que les résidus, différence entre les points expérimentaux et la droite de régression. Commenter.

4. Donner la concentration en glucose de la solution inconnue.

Partie 2 : Ecarts normalisés et compatibilité du résultat

On indique que les angles déviation sont mesurés à +/- 0,5 °.

1. Faire afficher la droite de régression ainsi que les résidus, différence entre les points expérimentaux et la droite de régression en tenant compte des barres d'erreur. Commenter.

2. Faire calculer puis afficher les écarts normalisés, entre la droite modèle et les points expérimentaux. Conclure.

Partie 3 : Méthode Monte-Carlo, Ecarts normalisés et compatibilité du résultat

On indique, par ailleurs, que les concentrations en masse en glucose sont données à +/- 1 g/L. Par ailleurs, la valeur théorique de la concentration en masse dans la solution S est de : 25 g/L.

3. En appliquant la méthode Monte-Carlo, réalisez N régressions linéaires à partir des données expérimentales.
4. En déduire la concentration en masse en glucose dans la solution S et lui associer une incertitude. Faire afficher les résultats.
5. Faire afficher l'écart normalisé et conclure.