

3.3 Incertitudes-type composées quelconques

Seules les deux formules précédentes sont à connaître, pour tous les autres cas, nous allons revenir à la définition des incertitudes puis, à l'aide d'une simulation informatique comportant une part d'aléatoire, calculer l'incertitude-type.

Définition. Un algorithme utilisant la variabilité d'une mesure pour simuler un calcul d'incertitude fait parti des **algorithmes de type Monte-Carlo**.

| *Remarque :* Les algorithmes de Monte-Carlo sont nombreux et ont de nombreuses applications. Leur point commun est qu'ils sont basés sur un processus aléatoire simulé informatiquement.

4 La régression linéaire

4.1 Principe

Définition. Prenons des listes de mesures x et y avec leurs incertitudes. La **régression linéaire** est une opération mathématique qui consiste à trouver les meilleurs coefficients a et b tels que $ax_i + b$ soient les plus proches en moyenne des points de mesures y_i .

| *Remarque :* En toute rigueur, il faudrait parler de régression affine et non pas linéaire.

Une régression linéaire permet de trouver la « meilleure droite » modélisant le mieux le comportement de ces points. Mathématiquement, on peut optimiser par un calcul dit « des moindres carrés » ce procédé.

🚫🚫🚫 **Attention !** Le coefficient r^2 n'a aucun intérêt pour valider un modèle physique ou pour estimer des incertitudes-type. Pour vous convaincre que le coefficient r^2 ne permet en rien de juger une régression linéaire, on peut regarder la vidéo suivante (dans laquelle le coefficient « cor » correspond au r) :

https://youtu.be/It4UA75z_KQ

4.2 Quand utiliser une régression linéaire ?

Supposons que l'on cherche à vérifier un modèle $y = ax + b$ et que l'on ait, après expérience, un ensemble de valeurs expérimentales $\{x_i\}$ et $\{y_i\}$ qui possèdent chacun une certaine variabilité.

La régression linéaire simple permet, à partir de l'ensemble des points expérimentaux, de trouver une valeur de a et une valeur de b . Pour estimer l'incertitude-type de ces paramètres, il faut réaliser des ensembles de nouvelles mesures $\{x_i, y_i\}$ puis réaliser une nouvelle régression linéaire. En réalisant un grand nombre de fois cette opération, on obtiendra, en prenant les écarts-types, les valeurs des incertitudes-types sur les paramètres. Bien souvent, un tel procédé est bien trop long et donc peu pratique.

Il faut donc mieux, lorsque c'est possible, éviter le traitement des incertitudes par régression linéaire.

- ▷ Lors d'une expérience identique réalisée par plusieurs expérimentateurs conduisant à une régression linéaire par chacun des expérimentateurs, l'incertitude-type finale est uniquement évaluée par l'étude de la variabilité des résultats de la régression linéaire. Il n'est pas nécessaire d'étudier plus en détail la régression linéaire. Dans ce cas, la régression linéaire est une des étapes du protocole de mesure, et sa variabilité est intégrée lors du calcul de variabilité final.
- ▷ Si le modèle recherché implique $b = 0$, dans ce cas, on cherche uniquement à estimer a . On peut donc calculer un grand nombre de valeur de a par la relation y_i/x_i puis réaliser un traitement statistique sur ces valeurs. Le grand intérêt est que, dans ce cas, les N points de mesures conduisent à N valeurs de a ayant une variabilité claire. Dans le cas de la régression linéaire, ces N valeurs conduisent à une unique valeur de a dont on ne peut pas directement étudier la variabilité.

Dans tous les autres cas, on peut estimer la variabilité de a et b par une simulation Monte-Carlo.

Remarque : Dans ce second cas, si le calcul de l'écart normalisé final entre la mesure de a et la valeur attendue a a une valeur plus grande que 2, c'est que le modèle linéaire n'est peut être pas optimal et qu'un modèle avec un b non nul serait préférable.

4.3 Application par méthode Monte-Carlo

On commence par réaliser une régression linéaire sans incertitudes sur les valeurs mesurées. La pente et l'ordonnée à l'origine obtenues sont le résultat numérique de la mesure.

Pour estimer leurs incertitudes-type, il faut de plus estimer **pour chaque point de mesure** la valeur de son incertitude-type. Ensuite, à l'aide d'une simulation Monte-Carlo, on utilise cette variabilité individuelle pour générer un grand nombre de distributions de points. Pour chacune de ces distributions, on réalise une régression linéaire ce qui conduira au final à un grand nombre de valeurs de a et de b . Il suffit ensuite de réaliser une étude statistique de ces données pour en déduire leur incertitude-type.

Pour s'assurer qu'une régression linéaire est correcte, on tracera **systématiquement** sur un graphique les données mesurées ainsi que la droite de la régression linéaire. Le modèle sera validé si, à l'œil, les points de mesure sont bien alignés et que la droite passe le plus proche de tous les points possibles, en incluant leurs incertitudes-type.

Remarque : On rappelle que l'incertitude-type est une estimation de la variabilité de la mesure. Ainsi, il est naturel que les points expérimentaux soient éloignés de la valeur de la modélisation de quelques incertitudes-type.

Si l'on constate par exemple que les points ressemblent plus à une parabole qu'à une droite, la régression linéaire ne sera pas l'outil approprié.

5 Annexe 1 : Méthode Monte-Carlo pour estimer des incertitudes-types

Les codes présentés dans cette partie sont réalisés sous python. Ce n'est absolument pas une obligation, les simulations peuvent être réalisées avec n'importe quel tableur pouvant réaliser un tirage de nombres aléatoires.

5.1 Incertitude-type composée

5.1.1 Principe

Supposons que l'on cherche à estimer une grandeur y donnée par $y = f(x_1, x_2, \dots)$ avec les x_i des données résultants d'une mesure et f une fonction connue. Chaque x_i est caractérisé par sa valeur et son incertitude-type.

La valeur de y est donnée par l'application de la formule. Pour estimer l'incertitude-type, il faut remonter à la variabilité de y , qui est elle-même une conséquence de la variabilité des x_i .

Pour cela, il faut :

- ▷ Fixer un nombre N de simulations à réaliser ;
- ▷ Pour k entre 1 et N , réaliser
 - ▷ un tirage aléatoire pour chaque x_1 ;
 - ▷ utiliser les valeurs de ce tirage et la fonction f pour calculer une valeur y_k ;
 - ▷ sauvegarder cette valeur y_k ;
- ▷ l'incertitude-type de y est l'écart-type de la distribution des y_k . La moyenne des y_k permet de retrouver la valeur y .

Le choix de la distribution de probabilité de chaque x_i dépend de plusieurs facteurs expérimentaux. La modélisation de cette distribution peut être délicate. Par exemple, il n'est pas du tout naturel de prendre systématiquement une distribution gaussienne.

En règle général en CPGE, les x_i sont mesurés avec une précision. En dessous de cette précision, l'étudiant est incapable d'accéder à une information sur la distribution de probabilité. On privilégie donc la **distribution uniforme de probabilité**, cohérente avec l'expérience pratique de l'étudiant. Toute autre forme de distribution doit pouvoir être justifiée et argumentée par l'étudiant.

🚫🚫🚫 **Attention !** Il ne faut pas oublier qu'en notant Δ la précision, l'incertitude-type qui caractérise la variabilité vaut $u = \Delta/\sqrt{3}$.

5.1.2 Exemples

À l'aide de python, on a simulé à l'aide d'une méthode Monte-Carlo les incertitudes composées de :

- ▷ une fréquence connaissant la période ;
- ▷ une longueur entre deux points connaissant leurs positions ;
- ▷ une célérité connaissant une fréquence et une longueur d'onde ;
- ▷ une célérité connaissant la période de l'onde et la position de deux nœuds séparant 10 longueurs d'ondes ;
- ▷ une distance focale connaissant les positions de l'objet et de l'image.

Tous ces exemples sont regroupés dans un [Google Colab disponible en cliquant ici](#).

5.2 Régression linéaire

5.2.1 Principe

En physique-chimie, nous n'avons jamais des séries parfaites de nombres, chaque mesure a une incertitude. Supposons que nous ayons réalisé m mesures de couples (x_i, y_i) , chacun avec une incertitude-type $(u(x_i), u(y_i))$. Ces incertitudes-types peuvent être différentes pour chaque mesure.

Pour en tenir compte, on réalise une simulation Monte-Carlo en réalisant un très grand nombre de régression linéaire sur une série de points sans incertitudes. Pour obtenir ces points, on génère aléatoirement pour chaque point mesuré une valeur à l'aide des incertitudes-type expérimentales. Conformément aux préconisations du GUM, si nous n'avons pas d'information pour savoir de quelle façon générer les valeurs, on choisit une distribution de probabilité uniforme (cf. section précédente). Attention, on rappelle que pour une distribution uniforme entre $x - \Delta$ et $x + \Delta$, on a la relation $\Delta = u(x) \times \sqrt{3}$.

Les valeurs finales de la pente et de l'ordonnée à l'origine sont les moyennes de toutes leurs valeurs, et leurs incertitudes-types sont les écarts-types de ces deux ensembles de valeurs.

On note m le nombre de points de mesures. Pour la simulation Monte-Carlo, il faut réaliser le programme ci-dessous décrit en pseudo-code :

- ▷ réaliser une régression linéaire unique pour estimer la pente et l'ordonnée à l'origine ;
- ▷ fixer un nombre N très grand ;
- ▷ créer deux listes vides pour stocker les pentes et les ordonnées à l'origine des régressions ;
- ▷ pour chaque i compris entre 1 et N , réaliser
 - ▷ pour chaque j compris entre 1 et m , réaliser un tirage aléatoire d'une valeur de Y_j donnée par une loi de probabilité uniforme entre $y_j - \sqrt{3} \times u(y_j)$ et $y_j + \sqrt{3} \times u(y_j)$;
 - ▷ pour chaque j compris entre 1 et m , réaliser un tirage aléatoire d'une valeur de X_j donnée par une loi de probabilité uniforme entre $x_j - \sqrt{3} \times u(x_j)$ et $x_j + \sqrt{3} \times u(x_j)$;
 - ▷ réaliser une régression linéaire sur cet ensemble (X_j, Y_j) puis ajouter dans les listes la pente et l'ordonnée à l'origine de cette régression.
- ▷ calculer pour les écarts-types des deux listes des pentes et des ordonnées à l'origine pour obtenir les incertitudes-types ;
- ▷ tracer sur un graphique la droite obtenue avec la pente moyenne et l'ordonnée à l'origine moyenne ;
- ▷ superposer sur le même graphique les points de mesures en indiquant leurs incertitudes-type sous la forme de barre d'erreur.

5.2.2 Exemple

Exercice rédigé avec Pierre Jamonneau.

Lorsqu'un métal est éclairé par un rayonnement ultraviolet ou proche ultraviolet, des électrons sont extraits du métal : c'est l'effet photoélectrique. On considère une plaque de baryum éclairée par une lampe spectrale au mercure. Un système de filtres permet de sélectionner une longueur d'onde particulière λ émise par la lampe. Un dispositif expérimental indépendant permet de mesurer l'énergie cinétique \mathcal{E}_c des électrons extraits de la plaque de baryum.

On récapitule ci-dessous les valeurs mesurées. L'incertitude-type de l'énergie cinétique est de 0.05 eV.

\mathcal{E}_c (eV)	2.40	1.69	0.91	0.57	0.35
ν ($\times 10^{14}$ Hz)	11.825	10.111	8.210	7.4129	6.8838
$u(\nu)$ ($\times 10^{12}$ Hz)	2.3	1.7	1.1	0.91	0.79

L'énergie cinétique des électrons extraits d'un métal éclairé par une onde électromagnétique de fréquence ν s'écrit :

$$\mathcal{E}_c = h(\nu - \nu_s) \quad (5.1)$$

où ν_s est une constante appelée fréquence seuil.

Le code python de la régression linéaire est disponible dans [un Google Colab en cliquant ici](#).

Par régression linéaire, on obtient $h = (6.63 \pm 0.20) \times 10^{-34}$ J · s et $\nu_s = (6.03 \pm 0.10) \times 10^{14}$ Hz.

Il est à noter que la fréquence de seuil doit être calculée en cours de simulation Monte-Carlo, et non pas à l'aide des valeurs finales de pente et d'ordonnée à l'origine moyenne. En effet, ce calcul en cours de simulation permet la prise en compte des corrélations entre la pente et l'ordonnée à l'origine. Dans cette expérience, on constate un facteur 3 entre les deux possibilités.