Incertitudes

Plan du cours

Même en s'appliquant du mieux possible, réaliser une même expérience dans les mêmes conditions conduira presque toujours à des résultats différents : on parle de **variabilité** des résultats de l'expérience. La première cause de variabilité est bien souvent l'expérimentateur lui-même, qui ne peut reproduire les mêmes gestes rigoureusement à l'identique. L'environnement et les instruments de mesure sont d'autres causes importantes.

L'**incertitude** est un nombre qui quantifie la variabilité des résultats obtenus lors de différentes réalisations d'une même expérience dans les mêmes conditions.

En physique, on utilise conventionnellement l'**incertitude-type**, c'est-à-dire reliée à un écart-type.

Attention ! Ne pas confondre incertitude et erreur dans la réalisation du protocole : ce n'est pas parce que les résultats varient d'une réalisation à l'autre que certaines mesures sont correctes et d'autres fausses. Ainsi, sauf si vous avez de bonnes raisons de croire que vous vous êtes trompé dans la réalisation de l'expérience, il n'y a aucune raison d'éliminer une mesure « parce que sa valeur est différente, donc fausse ».

Le résultat final d'une expérience de mesure d'une grandeur x doit nécessairement présenter la valeur expérimentale X_{exp} **et** l'incertitude-type $u(X_{\text{exp}})$ associée.

Convention : l'incertitude-type s'exprime avec deux chiffres significatifs, et la position du deuxième chiffre donne celle du dernier chiffre significatif à garder sur la valeur mesurée.

$$\begin{aligned} X_{\text{exp}} &= 17{,}306\ \text{cm} \\ u(X_{\text{exp}}) &= 0{,}287\ \text{cm} \end{aligned} \quad \text{devient} \quad \begin{cases} x = 17{,}31\ \text{cm} \\ u(x) = 0{,}29\ \text{cm} \end{cases}$$

I - Modélisation probabiliste d'un processus de mesure

Formellement, la mesure d'une grandeur physique se modélise par le tirage d'une variable aléatoire x, dont on obtient un résultat différent à chaque réalisation de l'expérience (= chaque tirage). « Mesurer la grandeur » revient à estimer du mieux possible l'espérance $\mathbb{E}(x)$ et l'écart-type σ_x de la variable aléatoire, appelées dans ce contexte **valeur expérimentale** et **incertitude-type**.

La loi de probabilité de la variable aléatoire x est (et restera !) inconnue : on ne peut que la modéliser. Les deux modèles que nous utiliserons sont la **loi normale** et la **loi uniforme**, représentées figure 1.

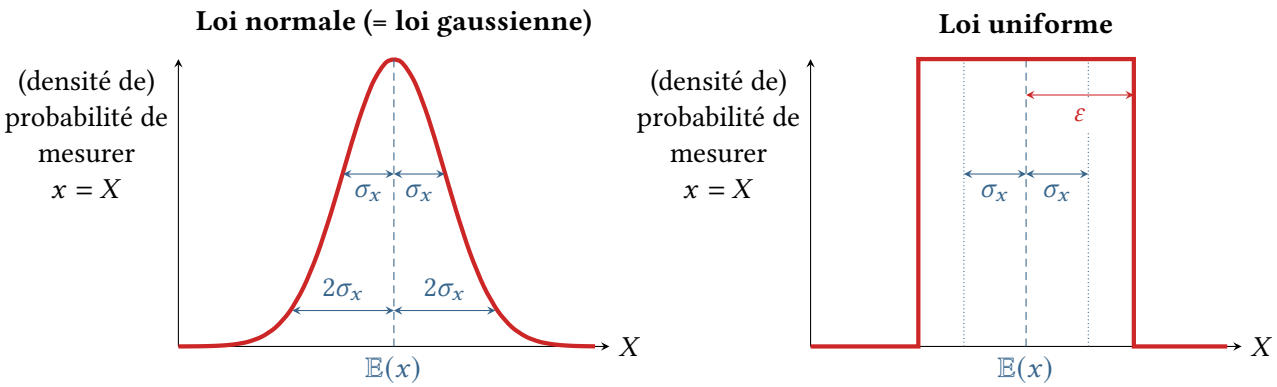

Figure 1 – Lois de probabilité usuelles.

- Pour une loi normale, on montre que la probabilité qu'une mesure tombe dans l'intervalle $[\mathbb{E}(x)-\sigma_x, \mathbb{E}(x)+\sigma_x]$ est de 68 % : on parle d'**intervalle de confiance à 68 %**. De la même façon, $[\mathbb{E}(x)-2\sigma_x, \mathbb{E}(x)+2\sigma_x]$ est l'intervalle de confiance à 95 %.
- Dans le cas d'une loi uniforme, l'écart-type σ_x est relié à la demi-largeur ε par

$$\sigma_x = \frac{\varepsilon}{\sqrt{3}}.$$

Pour cette loi, $[\mathbb{E}(x)-\sigma_x, \mathbb{E}(x)+\sigma_x]$ est l'intervalle de confiance à 57 % ... et bien sûr $[\mathbb{E}(x)-\varepsilon, \mathbb{E}(x)+\varepsilon]$ l'intervalle de confiance à 100 % !

II - Estimer l'incertitude sur une grandeur mesurée

II.A - La grandeur fluctue : estimation de type A

La situation la plus simple à analyser est celle où la variabilité est directement visible sur les résultats expérimentaux. On parle alors d'une **estimation d'incertitude de type A**.

Quand on dispose d'une série de mesures $X_1, ..., X_N$, la valeur expérimentale est la moyenne

$$X_{\text{exp}} = \langle X_n \rangle = \frac{1}{N}\sum_{n=1}^{N} X_n$$

et l'incertitude-type sur cette valeur est reliée à l'écart-type σ de la série de mesure par

$$u(X_{\text{exp}}) = \frac{\sigma}{\sqrt{N}}$$

Pour approfondir : Ces résultats s'établissent de manière mathématiquement rigoureuse dans le cas où la distribution de probabilité de la grandeur mesurée suit une loi normale. La distribution de probabilité exacte étant de toute façon inconnue et inaccessible, on les utilise dans tous les cas : on parle d'*approximation normale*, dont on peut montrer qu'elle est de mieux en mieux vérifiée que le nombre de mesures réalisées est élevé. Ce résultat très utilisé en statistiques porte le nom de *théorème central limite*. ■

Il existe plusieurs conventions de définition pour l'écart-type. Celle utilisée pour l'incertitude-type est

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\Big(X_n - \langle X_n\rangle\Big)^2} \qquad \text{où } \langle X_n\rangle \text{ désigne la moyenne.}$$

En mathématiques, l'écart-type est généralement défini avec une division par N plutôt que par $N-1$, mais on peut qualitativement comprendre qu'en physique, l'interprétation comme une incertitude exige de diviser par $N-1$. En effet, si on ne réalisait qu'une seule mesure ($N = 1$), diviser par N conduirait à une incertitude-type $u(x) = 0$ ... alors que le résultat n'est absolument pas certain ! Au contraire, diviser par $N-1$ donne une incertitude-type non définie (0 divisé par 0), ce qui est conforme au sens physique. L'argument optionnel `ddof=1` de la fonction `np.std` indique cette division par $N-1$.

Un code clés en main

```python
import numpy as np

x = np.array([ ... ]) # à compléter

x_exp = np.mean(x)
u_x_exp = np.std(x,ddof=1)/np.sqrt(len(x))

print('x =', x_exp)
print('u(x) =', u_x_exp)
```

II.B - La grandeur ne fluctue pas : estimation de type B

Il est des situations pour lesquelles reproduire l'expérience à l'identique ne permet pas d'observer de variabilité dans les résultats, ou bien parce que la grandeur elle-même ne varie pas dans les conditions de l'expérience (p.ex. largeur d'une table) ou bien parce que sa variabilité est inférieure à la résolution de l'instrument de mesure utilisé (p.ex. graduation d'une règle ou d'un rapporteur). Une situation analogue est celle où l'expérience ne peut être réalisée qu'une seule fois. Dans tous les cas, on ne dispose que d'un seul et unique résultat ... mais cela ne veut absolument pas dire qu'il est certain ! On parle dans ce contexte d'**incertitude de type B**.

Il faut alors estimer un intervalle dans lequel on est (raisonnablement) certain que le résultat de la mesure se trouve, noté sous la forme $[x_0 - \varepsilon, x_0 + \varepsilon]$, dont x_0 est appelé le centre et ε la **demi-étendue**. La demi-étendue est estimée ou bien « avec bon sens », ou bien en se reportant à la notice de l'instrument de mesure utilisé. On peut alors utiliser les résultats pour une loi uniforme,

Pour un résultat ne fluctuant pas, compris avec « certitude » dans l'intervalle $[x_0 - \varepsilon, x_0 + \varepsilon]$,

$$X_{\text{exp}} = x_0 \qquad \text{et} \qquad u(X_{\text{exp}}) = \frac{\varepsilon}{\sqrt{3}}$$

Un exemple pour comprendre :

Imaginons avoir l'idée farfelue de mesurer la longueur d'une carte bancaire avec la règle jaune du tableau.

On constate que la longueur de la carte est comprise avec certitude dans l'intervalle [8 cm; 9 cm]. On en déduit que la longueur est de 8,5 cm, avec une incertitude-type $0{,}5/\sqrt{3} = 0{,}29$ cm.

> ***Remarque :*** *J'ai ici considéré le cas le plus simple où la demi-étendue ε correspond à la moitié d'une graduation, mais ce n'est pas une règle générale : tout dépend de la situation et de l'appréciation de l'expérimentateur. Par exemple, mesurer avec la même règle jaune la largeur d'une feuille A4 donnerait 21 cm « presque tout pile » : compte tenu de l'observation visuelle de la règle, on peut par exemple prendre comme intervalle* [20,8 cm; 21,2 cm].

III - Estimer l'incertitude sur une grandeur calculée

Sauf exception, l'incertitude sur une grandeur calculée ne se devine pas de façon immédiate en raison du « principe de compensation des erreurs » : une fluctuation sur une valeur intervenant dans le calcul peut être compensée par une autre, au moins partiellement. L'incertitude est appelée **incertitude composée**, et on parle de **composition des incertitudes** sur les différentes grandeurs.

Attention ! Avant de se lancer dans un calcul d'incertitude-type composée, il faut toujours commencer par comparer entre elles les incertitudes relatives $u(y)/y$, $u(z)/z$, etc, des différentes grandeurs d'entrée et ne considérer que celles qui sont prépondérantes. Très souvent, une de ces grandeurs possède une incertitude relative très supérieure aux autres, ce qui évite un calcul fastidieux.

III.A - Formules de composition des incertitudes

Multiplication par une constante	$x = \lambda y \quad (\lambda \in \mathbb{R})$	$u(x) = \lvert\lambda\rvert\, u(y)$
Somme ou différence	$x = y + z \quad \underline{\text{ou}} \quad x = y - z$	$u(x) = \sqrt{u(y)^2 + u(z)^2}$
Produit ou quotient	$x = y \times z \quad \underline{\text{ou}} \quad x = \dfrac{y}{z}$	$\dfrac{u(x)}{x} = \sqrt{\left(\dfrac{u(y)}{y}\right)^2 + \left(\dfrac{u(z)}{z}\right)^2}$
Puissances et racines	$x = y^p \quad (p \in \mathbb{R})$	$\dfrac{u(x)}{x} = \lvert p\rvert \dfrac{u(y)}{y}$

Pour approfondir : Ces formules de composition se démontrent par des raisonnements probabilistes. En toute rigueur, elles ne sont valables que dans la limite d'un nombre infini de mesures suivant une loi de probabilité gaussienne. Il existe également des formules plus générales, valables pour n'importe quelle fonction $x = f(y, z, ...)$, mais elles conduisent à des calculs souvent complexes et peu parlants, si bien qu'elles ne figurent pas aux programmes de CPGE.

Contrairement à ce que l'on pourrait croire, la dernière ligne y^p ne peut pas se déduire de l'avant-dernière $y \times z$. Pour que ces relations s'appliquent, les deux variables y et z doivent être indépendantes l'une de l'autre ... ce qui n'est évidement pas le cas de y avec lui-même. Ce faisant, il y a une compensation partielle des incertitudes sur y et z : de manière très schématique, si une fluctuation rend y un peu plus grand, il se peut que z soit rendu un peu plus petit, ce qui donne une compensation dans le calcul de x. Un tel effet est bien sûr impossible dans le calcul de y^p. ■

III.B - Estimation numérique par simulation Monte Carlo

- **Principe**

On cherche à estimer une grandeur y reliée à un ensemble de grandeurs expérimentales $x_1, x_2, ..., x_K$ par une relation écrite formellement

$$y = f(x_1, x_2, ..., x_K)\,,$$

avec f une fonction quelconque mais connue. Chaque grandeur expérimentale x_k ($1 \leq k \leq K$) est caractérisée par sa valeur et son incertitude-type $u(x_k)$. Pour estimer l'incertitude-type $u(y)$, on recourt à la simulation de $N \gg 1$ expériences.

- Pour chaque réalisation n de l'expérience,
  - → on tire aléatoirement une valeur $x_{k,n}$ pour chaque x_k, en utilisant une loi normale ou une loi uniforme selon la façon dont l'incertitude-type $u(x_k)$ est estimée ;
  - → et on en déduit la valeur correspondante y_n.
- L'incertitude-type composée $u(y)$ correspond à l'écart-type de la série des y_n simulés.

Pour approfondir : Pourquoi prendre $u(y) = \sigma$ et pas $u(y) = \sigma/\sqrt{N}$ comme précédemment ? Constatons d'emblée que dans le second cas plus on simulerait d'expériences, plus l'incertitude-type diminuerait ... alors que l'on dispose toujours des mêmes informations sur le système expérimental. Réduire l'incertitude sans affiner la connaissance du système serait contradictoire. L'explication vient du fait que la simulation permet d'estimer l'incertitude sur *une seule* mesure de y, celle qui a été réalisée « pour de vrai » dans l'expérience, et non pas sur une moyenne des résultats de *plusieurs* mesures. ■

En pratique, les fonctionnalités des tableaux `numpy` permettent de simuler les `N` expériences sans faire appel à des boucles. En fonction de la manière dont l'incertitude-type $u(x_k)$ a été estimée, on choisira d'utiliser pour les tirages aléatoires une loi de probabilité gaussienne (la valeur centrale est plus probable que les autres) ou une loi de probabilité uniforme (toutes les valeurs de l'intervalle sont équiprobables). On utilisera les fonctions Python suivantes :

- pour le tirage de `N` valeurs aléatoires suivant une loi uniforme dans un intervalle [`mini`, `maxi`] : `np.random.uniform(mini,maxi,N)` ;
- pour le tirage de `N` valeurs aléatoires suivant une loi normale de moyenne `moy` et d'écart-type `sigma` : `np.random.normal(moy,sigma,N)`.

> ***Remarque culturelle :*** *Ces méthodes par tirage aléatoire étaient (sont ?) utilisées par les joueurs de casino pour estimer leurs gains potentiels, d'où leur nom générique de « méthodes de Monte Carlo ».*

Un code clés en main

```python
import numpy as np

x1 = ...
u_x1 = ... # on a l'incertitude (c'est un exemple !)

x2 = ...
eps_x2 = ... # on a la demi-étendue (c'est un exemple !)

y_exp = f(x1,x2) # remplacer f par son expression !

N = 1e4 # nombre de simulations
x1_sim = np.random.normal(x1, u_x1, N)
x2_sim = np.random.uniform(x2-eps_x2, x2+eps_x2, N)
y_sim = f(x1_sim, x2_sim) # calcul terme à terme car np.array
    # remplacer f par son expression !

u_y = np.std(y_sim,ddof=1)
# il est inutile de calculer np.mean(y_sim) !

print('y =', y_exp)
print('u(y) =', u_y)
```

IV - Comparaison entre deux valeurs

Les deux valeurs comparées peuvent être issues de deux expériences différentes (deux protocoles ou deux binômes), ou bien l'une d'elles être une valeur attendue donnée par un fabricant ou prévue par un modèle théorique.

L'idée est qualitativement simple : pour que la valeur mesurée soit compatible avec la valeur attendue, il ne faut pas qu'elle en soit « trop » éloignée. Comme la dispersion des valeurs expérimentales est quantifiée par l'incertitude-type, l'écart entre les valeurs mesurée et attendue doit être comparée à cette incertitude-type.

IV.A - Comparaison entre deux valeurs connues avec incertitude, z-score

On appelle **écart normalisé** ou ***z*-score** entre deux valeurs x_1 et x_2,
connues avec une incertitude type $u(x_1)$ et $u(x_2)$,

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

Par convention, les deux valeurs x_1 et x_2 sont dites **compatibles** si

$$z \leq 2\,.$$

Interprétation graphique : On simule 10 000 réalisations de deux expériences, voir figure 2, représentées sous forme des deux histogrammes. Si le z-score est suffisamment faible, les histogrammes issus des deux expériences se recouvrent « beaucoup », alors qu'ils ne se recouvrent que « très peu » lorsque le z-score est élevé.

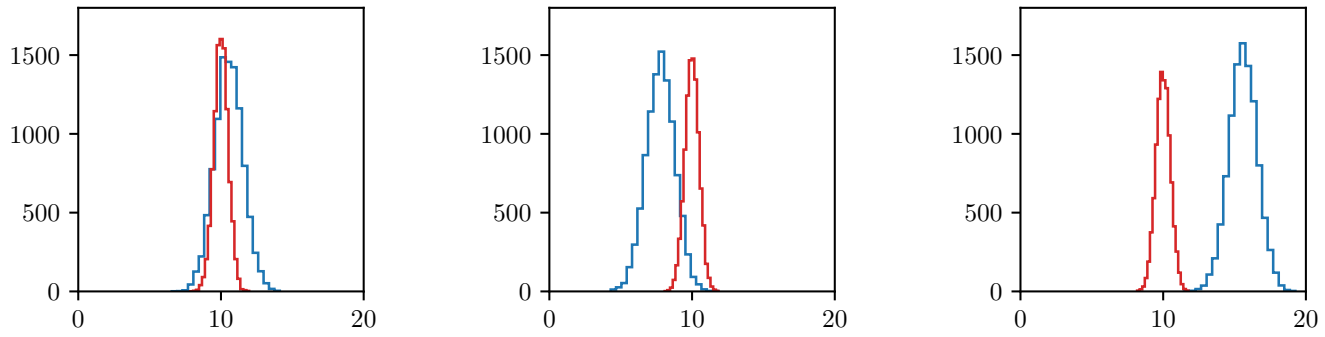

Figure 2 – Critère de compatibilité entre deux valeurs sur lesquelles l'incertitude-type est connue. De gauche à droite : $z = 0{,}5$ (les résultats des deux expériences sont considérées compatibles), $z = 2$ (limite de compatibilité) et $z = 5$ (résultats incompatibles). Ce nombre de réalisations est bien sûr inatteignable dans une expérience réelle en CPGE, mais permet de bien visualiser les choses.

Interprétation mathématique : Les formules de composition des incertitudes données au paragraphe III.A indiquent que le terme au dénominateur du z-score peut s'écrire

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

c'est-à-dire qu'il s'interprète comme l'incertitude-type de la grandeur calculée $x_1 - x_2$. Ainsi, le z-score compare la valeur de $x_1 - x_2$ (calculée par exemple par moyennage) à son incertitude-type.

Si les valeurs de x_1 et x_2 sont compatibles, alors $x_1 - x_2$ devrait être nul ... ou en tout cas « pas très grand », l'égalité stricte étant inaccessible à cause de la variabilité intrinsèque aux expériences. Cette variabilité étant quantifiée par l'incertitude-type $u(x_1 - x_2)$, on comprend que le critère de compatibilité repose sur un z-score petit. Le seuil exact de compatibilité $z \leq 2$ est choisi par convention, essentiellement pour des raisons historiques : en pratique, les mesures ne sont pas « beaucoup plus compatibles », c'est-à-dire que les histogrammes ne se chevauchent pas beaucoup plus, si $z = 1{,}9$ que si $z = 2{,}1$.

IV.B - Comparaison entre une valeur expérimentale et une valeur connue sans incertitude

Il est fréquent que la valeur attendue $x^\star$ soit donnée sans incertitude ou que celle-ci soit négligeable, comme par exemple pour une grandeur tabulée. Dans ce cas, le critère sur le z-score devient

$$z = \frac{\left|x^\star - x_{\text{exp}}\right|}{u(x_{\text{exp}})} \leq 2$$

ce qui peut se reformuler de façon plus lisible sous la forme ci-dessous :

Par convention, la valeur expérimentale x_{exp} et la valeur attendue $x^\star$ sont dites compatibles si

$$\left|x^\star - x_{\text{exp}}\right| \leq 2u(x_{\text{exp}}) \quad \Longleftrightarrow \quad x^\star \in [x_{\text{exp}} - 2\,u(x_{\text{exp}});\ x_{\text{exp}} + 2\,u(x_{\text{exp}})]$$

Interprétation graphique : voir figure 3. Qualitativement, la valeur attendue est compatible avec les résultats expérimentaux si elle se trouve « à l'intérieur » de l'histogramme des résultats obtenus.

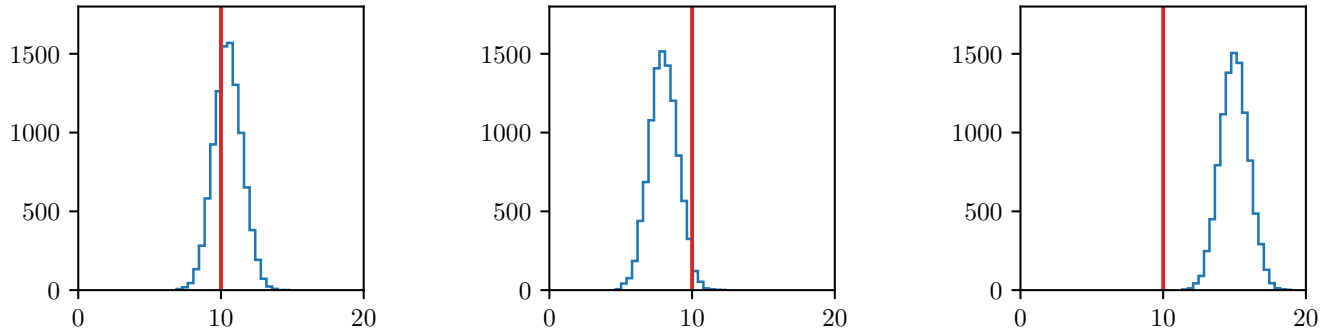

Figure 3 – Critère de compatibilité entre deux valeurs dont l'une est connue exactement. On simule 10 000 réalisations d'une même expérience, représentées sous forme de l'histogramme, à comparer à la valeur attendue représentée par le trait plein. De gauche à droite : $z = 0{,}5$ (valeurs considérées compatibles), $z = 2$ (limite de compatibilité) et $z = 5$ (valeurs incompatibles). Ce nombre de réalisations est bien sûr innatteignable dans une expérience réelle en CPGE, mais permet de bien visualiser les choses.

V - Validation d'un modèle théorique

V.A - Principe général

Valider un modèle recouvre deux aspects : d'une part, vérifier que les mesures expérimentales sont cohérentes avec la forme fonctionnelle attendue, et d'autre part vérifier que les valeurs des coefficients qui s'en déduisent sont compatibles avec les valeurs attendues.

Au niveau CPGE, la validation passe nécessairement par une confirmation graphique sur un tracé de la pertinence du modèle. La droite étant la seule courbe reconnaissable à l'œil sans ambiguité, il faut toujours réécrire le modèle testé sous forme linéaire ou affine en introduisant des variables auxilliaires si besoin. Les premières étapes sont dans tous les cas les suivantes :

❶ Réaliser N expériences pour plusieurs valeurs X_n ($1 \leq n \leq N$), et mesurer à chaque fois la valeur Y_n correspondante.

❷ Représenter graphiquement y en fonction de x : s'il ne s'agit notoirement pas d'une droite (p.ex. présence d'une courbure), il est inutile de poursuivre. Cette étape est indispensable : les calculs se calculeront toujours, mais cela ne garantit pas la pertinence du modèle, qui ne peut être testée à notre niveau *que* graphiquement.

V.B - Validation d'un modèle théorique linéaire $y = ax$

N'avoir qu'un seul coefficient à estimer permet une exploitation relativement simple pour vérifier la compatibilité de la valeur expérimentale a_{exp} à la valeur attendue $a^\star$. L'idée est de traiter chaque réalisation de l'expérience comme une mesure indépendante de la grandeur $a = y/x$. Comme on dispose de plusieurs mesures de la même grandeur, il suffit ensuite de dérouler la procédure « classique » d'une estimation de type A.

❸ Calculer la valeur A_n obtenue pour chaque réalisation de l'expérience, puis la valeur moyenne $a_{\text{exp}} = \langle A_n \rangle$.

❹ Estimer l'incertitude type $u(a_{\text{exp}}) = \sigma/\sqrt{N}$ comme décrit au paragraphe II.A.

❺ Comparer a_{exp} à la valeur attendue $a^\star$ par un calcul de z-score et conclure.

> ***Remarque :*** *La procédure proposée ici permet une estimation simple et rapide de l'incertitude $u(a_{exp})$, qui ne fait pas appel aux incertitudes $u(x), u(y)$ sur les valeurs mesurées. C'est pratique dans certains cas ... mais si on dispose de ces incertitudes, il est logique (et préférable !) d'en tenir compte pour estimer $u(a_{exp})$. Il faut alors procéder par une méthode Monte Carlo analogue à celle utilisée pour valider un modèle affine, décrite paragraphe V.C.*

Un code clés en main

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

x = np.array([ ... ])
y = np.array([ ... ])

a = y/x # calcul terme à terme car np.array
a_exp = np.mean(a)
u_a = np.std(a,ddof=1)/sqrt(len(a))

plt.figure()
plt.plot(x,y,'+')
plt.plot(x,a_exp*x,'-') # droite de régression
   # attention, a_exp et non pas a
plt.show()

print('a =', a_exp)
print('u(a) =', u_a)
```

V.C - Validation d'un modèle théorique affine $y = ax + b$

Les premières étapes sont les mêmes que précédemment, mais les coefficients a et b et leurs incertitudes-types ne peuvent pas être estimés aussi simplement.

- **Estimation des coefficients a_{exp} et b_{exp}**

❸ Réaliser une **régression linéaire** pour obtenir a_{exp} et b_{exp}.

Un algorithme de régression linéaire permet d'obtenir la meilleure estimation de a et b, c'est-à-dire les valeurs qui donnent la droite passant au plus près des points expérimentaux (X_n, Y_n). Avec Python, on les obtient grâce à la fonction `np.polyfit` par une syntaxe de la forme `a_exp, b_exp = np.polyfit(x,y,1)`. Le troisième argument « `1` » indique à la fonction que l'on modélise la courbe par un polynôme d'ordre 1.

➽ **Pour approfondir :** Plus formellement, l'algorithme détermine les valeurs de a et b qui minimisent « une » distance entre la droite théorique $y = ax + b$ et les points expérimentaux. Toute la difficulté consiste à définir correctement la fonction distance à utiliser. La définition utilisée par `polyfit` est celle des moindres carrés, où la distance est simplement définie par

$$d = \sum_{n=1}^{N} \Big(Y_n - (aX_n + b) \Big)^2 .$$

D'autres définitions sont possibles, en particulier pour prendre en compte les incertitude sur les points expérimentaux, l'idée étant qu'un point de faible incertitude doit avoir plus de poids qu'un point à l'incertitude élevée. ■

- **Estimation des incertitudes $u(a_{\text{exp}})$ et $u(b_{\text{exp}})$**

Si l'on ne dispose pas des incertitudes $u(x), u(y)$ sur les points expérimentaux, il n'existe pas de méthode aussi simple qu'un calcul d'écart-type pour estimer les incertitudes sur a_{exp} et b_{exp}.

> ***Remarque :*** *Des formules analytiques existent, reposant sur l'écart entre les points expérimentaux et la droite ... et certaines hypothèses simplificatrices sur l'incertitude à l'origine de ces écarts.*

La seule méthode au programme concerne le cas où l'on dispose de l'incertitude $u(y) \neq 0$ sur une des deux grandeurs mesurées alors que l'incertitude $u(x) = 0$ sur l'autre grandeur est supposée nulle [1]. Elle s'appuie sur une simulation Monte Carlo de K séries de mesures et les régressions associées.

1. Cette hypothèse qui peut sembler grossière reste souvent pertinente, car il est fréquent que l'une des deux grandeurs soit connue avec une incertitude relative très supérieure à l'autre. C'est alors celle-ci qu'il faut placer en ordonnée.

❹ ▹ Simuler K séries de mesures différentes des valeurs $y_{k,1}, y_{k,2}, ..., y_{k,N}$ tenant compte des incertitudes.
▹ Pour chaque série de mesures simulée, réaliser une régression linéaire, qui donne des coefficients a_k et b_k différents à chaque simulation.
▹ Les incertitudes types $u(a_{\text{exp}})$ et $u(b_{\text{exp}})$ sont égales à l'écart-type des séries des a_k et b_k simulés.

❺ Comparer les valeurs a_{exp} et b_{exp} aux valeurs attendues $a^\star$ et $b^\star$ par un calcul de z-score et conclure.

Une validation graphique complémentaire est possible en vérifiant que la distance entre les points et la droite de régression est inférieure ou de l'ordre de $2u(y)$. Cela peut se faire visuellement en représentant des **barres d'incertitudes** grâce à la fonction `plt.errorbar(x,y,yerr=u_y)`.

Un code clés en main

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

x = np.array([ ... ])
y = np.array([ ... ])
u_y = ... # valeur unique ou liste

a_exp, b_exp = np.polyfit(x,y,1)

K = 10000 # nombre de simulations
a_sim = []
b_sim = []

for k in range(K):
    y_sim = np.random.normal(y,uy) # pas N car y est déjà un np.array
    a,b = np.polyfit(x,y_sim,1)
    a_sim.append(a)
    b_sim.append(b)

u_a_exp = np.std(a_sim)
u_b_exp = np.std(b_sim)

plt.figure()
plt.plot(x,y,'+')
plt.plot(x,a_exp*x+b_exp,'-') # droite de régression
    # attention, a_exp,b_exp et non pas a,b
plt.show()

print('a =', a_exp, 'u(a) =', u_a_exp)
print('b =', b_exp, 'u(b) =', u_b_exp)
# Inutile de calculer np.mean(a_sim) et np.mean(b_sim)
```