

Incertitudes

PCSI₂ 2023 – 2024

Introduction : *les notions et capacités identifiées ci-dessous couvrent les deux années de formation en classe préparatoire aux grandes écoles ; leur pleine maîtrise est donc un objectif de fin de seconde année. Elles sont communes aux enseignements de physique et de chimie et leur apprentissage s'effectue de manière coordonnée entre les enseignants.*

L'accent est mis sur la variabilité de la mesure d'une grandeur physique et sa caractérisation à l'aide de l'incertitude type. La comparaison entre deux valeurs mesurées d'une même grandeur physique est conduite au moyen de l'écart normalisé, l'objectif principal étant de développer l'esprit critique des étudiants en s'appuyant sur un critère quantitatif. Le même esprit prévaut dans l'analyse des résultats d'une régression linéaire qui ne saurait s'appuyer sur l'exploitation non raisonnée du coefficient de corrélation (R^2).

Le recours à la simulation vise à illustrer, sur la base de mesures expérimentales, différents effets de la variabilité de la mesure d'une grandeur physique dans les cas des incertitudes types composées et de la régression linéaire.

Notions et capacités exigibles (programme officiel) :

- Variabilité de la mesure d'une grandeur physique. Incertitude. Incertitude-type.
 - Identifier les incertitudes liées, par exemple, à l'opérateur, à l'environnement, aux instruments ou à la méthode de mesure.
 - Procéder à l'évaluation d'une incertitude-type par une approche statistique (évaluation de type A).
 - Procéder à l'évaluation d'une incertitude-type par une autre approche que statistique (évaluation de type B).
 - Associer un intervalle de confiance à l'écart-type dans l'hypothèse d'une distribution suivant la loi normale.
- Incertitudes-types composées.
 - Évaluer l'incertitude-type d'une grandeur s'exprimant en fonction d'autres grandeurs, dont les incertitudes-types sont connues, à l'aide d'une somme, d'une différence, d'un produit ou d'un quotient.
 - Comparer entre elles les différentes contributions lors de l'évaluation d'une incertitude-type composée.
 - ☒ Simuler, à l'aide d'un langage de programmation ou d'un tableur, un processus aléatoire permettant de caractériser la variabilité de la valeur d'une grandeur composée.
- Écriture du résultat d'une mesure.
 - Écrire, avec un nombre adapté de chiffres significatifs, le résultat d'une mesure.
- Comparaison de deux valeurs ; écart normalisé.
 - Comparer deux valeurs dont les incertitudes-types sont connues à l'aide de leur écart normalisé.
 - Analyser les causes d'une éventuelle incompatibilité entre le résultat d'une mesure et le résultat attendu par une modélisation.
- Régression linéaire.
 - Utiliser un logiciel de régression linéaire afin d'obtenir les valeurs des paramètres du modèle.
 - Analyser les résultats obtenus à l'aide d'une procédure de validation : analyse graphique intégrant les barres d'incertitude ou analyse des écarts normalisés.
 - ☒ Simuler, à l'aide d'un langage de programmation ou d'un tableur, un processus aléatoire de variation des valeurs expérimentales de l'une des grandeurs – simulation Monte-Carlo – pour évaluer l'incertitude sur les paramètres du modèle.

I Valeur mesurée et incertitude type (mesurage)

1. Variabilité de la mesure d'une grandeur physique

En science expérimentale, un mesurage est une procédure complexe qui se traduit par une variabilité de la valeur mesurée : une répétition de cette mesure conduit naturellement à une dispersion des valeurs observées. Cette variabilité fait partie du mesurage, **il ne s'agit pas d'une erreur de mesure.**

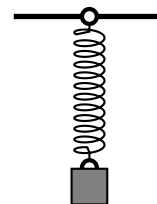
Mesurage

Le résultat d'un mesurage n'est pas une valeur numérique unique mais plutôt un ensemble de valeurs "raisonnablement attribuables" à la grandeur étudiée.



FIGURE 1 – Résultat de la mesure d'une grandeur X

Exemple : vous disposez d'un système " Masse - ressort vertical " qui constitue un oscillateur mécanique dont on cherche à mesurer la période des oscillations observées lors du mouvement vertical.



Il s'agit d'une manipulation dont vous devez systématiquement préciser (Cf annexe de TP) : **l'objectif, le protocole détaillé, les résultats obtenus et la conclusion déduite.**

Ecrivons cela directement dans un Notebook  Colab¹ pour exploiter la puissance de Python.



Notebook Colab



2. Meilleur estimateur et incertitude type associés à un mesurage

Considérons un ensemble $\{ x_1, \dots, x_N \}$ de N observations de la valeur d'une même grandeur physique X . Voici ce qu'on peut observer dans certains cas en les plaçant sur un axe.

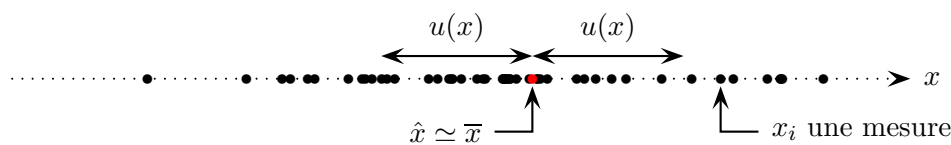


FIGURE 2 – Représentation d'un ensemble $\{ x_i \}$ de 50 observations

Méthode :

Pour effectuer des tracés sous python, on a besoin du module `matplotlib.pyplot`

```
import matplotlib.pyplot as plt
```

Définition

L'incertitude type, notée $u(x)$, caractérise la variabilité du processus de mesurage de la grandeur physique X autour du meilleur estimateur \hat{x} .

1. Il existe d'autres solutions que Colab de Google qui a pour principal avantage d'être utilisable sans installation particulière sur l'ordinateur ou la tablette.

$u(x)$ dépend de la nature du protocole expérimental mis en œuvre, du matériel utilisé, des conditions expérimentales et de l'habileté des expérimentateurs.

Sur la figure précédente, on constate que :

- Il y a plus de points à proximité de \hat{x} : la probabilité d'obtenir x_i dans cette zone est plus importante.
- L'incertitude type $u(x)$ constitue l'ordre de grandeur de la distance séparant deux valeurs mesurées prises au hasard.

Si N est assez grand, on considère que $\hat{x} = \bar{x}$ la valeur moyenne et $u(x)$ est l'écart type des valeurs x_i .

Lien avec les fonctions statistiques

On caractérisera le résultat d'une mesure par son meilleur estimateur $\hat{x} \simeq \bar{x}$ la valeur moyenne et son incertitude type égale à l'écart type $u(x)$ de la distribution $\{x_i\}$ avec

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{et} \quad u(x) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Remarque : le nombre N doit, en toute rigueur, être très élevé $N > 1000$.

Exemple : afin d'illustrer la variabilité sur $10T$, nous allons nous contenter de réaliser $n = 12$ mesures.

Méthode :

L'utilisation de tableaux dans python (requiert le module `numpy` : on tape `import numpy as np`) permet un calcul rapide de

- \bar{T} par la commande `T.mean()` et
- $u(T)$ par la commande `T.std(ddof=1)`

si T est un tableau de valeurs numériques (`dtype=float`).



3. Approche algorithmique, méthode de Monte Carlo

a. Principe

En CPGE le traitement des problèmes d'incertitude sera principalement réalisé à l'aide de **simulations Monte Carlo** qui se basent sur des procédés aléatoires.²

Méthode :

Si on connaît ou qu'on peut estimer l'incertitude type $u(x)$ associée à la mesure d'une grandeur X , on construit à l'aide de python le tableau de N valeurs par tirage aléatoire de valeurs $\{x_i\}$ avec $i = 1, \dots, N$ où $N \gg 1$ autour d'une valeur $\hat{x} \sim \bar{x}$ et avec un écart type $u(x)$.

Reste à déterminer l'incertitude type sur la grandeur mesurée.

b. Expériences avec variabilité observée (incertitude de type A)

Contexte et mise en application : lorsque la variabilité de la valeur x d'une grandeur physique X est effectivement observable, on réalise un grand nombre N de fois ce même processus.

Le meilleur estimateur correspond à la moyenne arithmétique des observations et l'incertitude type est égal à l'écart type.

Exemple : lors de la mesure de T on a travaillé avec une douzaine de points mais si on en prenait beaucoup plus on obtiendrait certainement une distribution "normale" d'allure "gaussienne".

2. Le nom de cette méthode fait allusion aux jeux de hasard pratiqués au casino de Monte Carlo.

Méthode :

La méthode `plt.hist(T)` permet d'afficher très simplement le nombre de fois où on obtient telle ou telle valeur de T sous la forme d'un histogramme.



Intervalle de confiance pour une loi normale : dans le cas où la valeur mesurée X peut être décrite comme une variable aléatoire distribuée selon une loi de probabilité normale (gaussienne), il est possible d'associer un intervalle de confiance à sa répartition autour de la valeur la plus probable.

Intervalle de confiance lié à une loi normale

Dans le cas d'une variable aléatoire x distribuée selon une loi normale d'espérance \hat{x} et d'écart type $u(x)$ (correspondant à l'incertitude type dans le cas où x est une grandeur mesurable) :

- l'intervalle $[\hat{x} - u(x), \hat{x} + u(x)]$ correspond à un niveau de confiance de 68,3 %
- l'intervalle $[\hat{x} - 2u(x), \hat{x} + 2u(x)]$ correspond à un niveau de confiance de 95,4 %
- l'intervalle $[\hat{x} - 3u(x), \hat{x} + 3u(x)]$ correspond à un niveau de confiance de 99,7 %

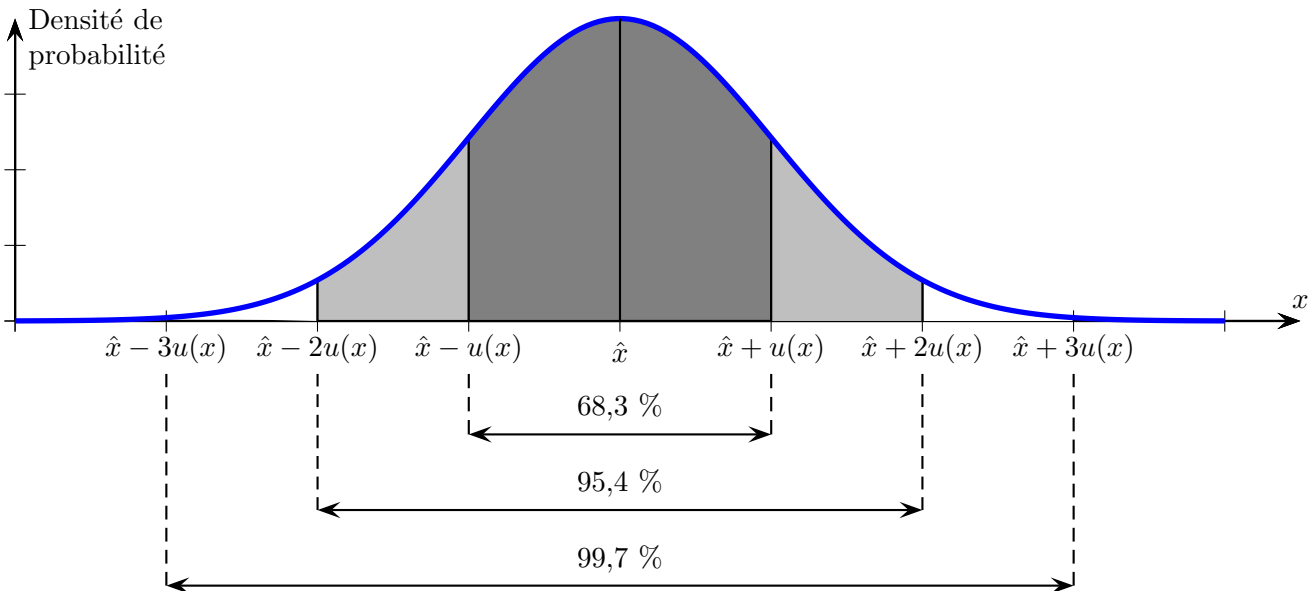


FIGURE 3 – Distribution normale d'une variable aléatoire

Certains points peuvent être très éloignés de la valeur moyenne \bar{x} : ce type d'événement est peu fréquent, mais pas improbable.

Méthode :

La méthode `np.random.randn(N)` génère un tableau de N valeurs réparties de façon gaussienne avec un écart type de 1.

En multipliant par $u(x)$ puis en ajoutant \bar{x} on simule le résultat d'un mesurage de N valeurs !



$$\hat{x} + u(x) * \underbrace{\text{np.random.randn}(N)}_{-4 \leq \dots \leq 4}$$

$$\underbrace{\phantom{\hat{x} + u(x) * \text{np.random.randn}(N)}}_{-4u(x) \leq \dots \leq 4u(x)}$$

$$\underbrace{\phantom{\hat{x} + u(x) * \text{np.random.randn}(N)}}_{\hat{x} - 4u(x) \leq \dots \leq \hat{x} + 4u(x)}$$

`x+ux*np.random.randn(N)`

Exemple : compléter le notebook pour déterminer T , la période de l'oscillateur. Visualiser la distribution et vérifier son écart type.

Incertitude type de la moyenne sur n mesurages : admettons qu'on ait effectué n mesures de la grandeur X , chacune possède une incertitude type $u(x)$ mais au final les fluctuations positives des observations compenseront, au moins en partie, les fluctuations négatives.

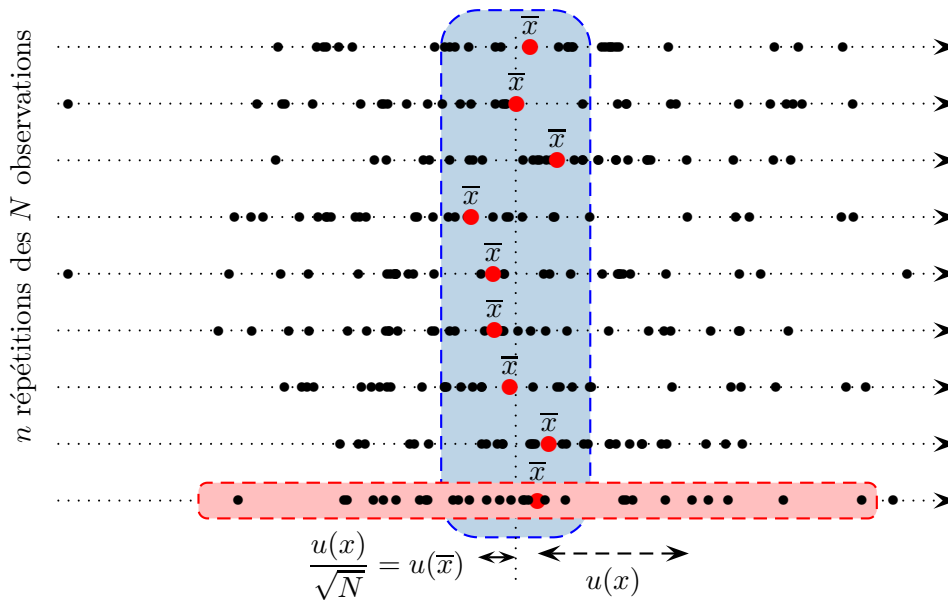


FIGURE 4 – Répétition de la série de N observations

Cela réduit de manière significative de la variabilité potentielle du résultat ! A la place de l'incertitude type $u(x)$ associée à un mesurage unique, on obtient une incertitude type $u(\bar{x}) < u(x)$ sur la mesure de la moyenne \bar{x} déduite de n observations. On peut en effet montrer que $u(\bar{x}) = \frac{u(x)}{\sqrt{N}}$ si $N \gg 1$.

Incertitude de type A

La répétition d'un mesurage permet une estimation plus précise de la valeur d'une grandeur physique X en diminuant la variabilité $u(x)$ associée à un mesurage unique.

Remarque : on pourrait demander aux n binômes de mesurer la période d'un seul oscillateur mécanique puis de calculer la moyenne, l'écart type serait plus faible.

c. Ecriture du résultat d'une mesure

Informations associées au résultat d'un mesurage

Le résultat de la mesure d'une grandeur physique X doit inclure le meilleur estimateur \hat{x} de la valeur x de cette grandeur et l'incertitude type associée $u(x)$.

On peut noter ce résultat sous l'une des formes suivantes :

$$X = \hat{x} \pm u(x) \quad \text{ou} \quad X = \{\hat{x}; u(x)\}$$

Il est important de ne pas oublier l'unité commune de ces deux paramètres \hat{x} et $u(x)$.

Méthode :

Procédure d'écriture standardisée :

- ① on arrondit l'incertitude type $u(x)$ à 1 ou 2 chiffres significatifs (maximum).
- ② on fait en sorte que le dernier chiffre du meilleur estimateur \hat{x} possède la même position (dans la mantisse ou en écriture décimale) que le dernier chiffre de l'incertitude type.
- ③ la méthode

$$\text{round}(x, n)$$

permet de garder n décimales du réel x

Exemples :

- Détermination d'une résistance électrique.

Grandeur physique mesurée	Valeur mesurée	Incertitude type
Résistance électrique R	4,20317 k Ω	5,236 Ω
Ecriture standardisée	$R = 4203,2 \pm 5,2 \Omega$ $\{ \hat{R} = 4203,2 \Omega ; u(R) = 5,2 \text{ k}\Omega \}$	

- Ecrire le résultat de la mesure de la période de l'oscillateur mécanique.

Grandeur physique mesurée	Valeur mesurée	Incertitude type
Période des oscillations	0.9439750095796157	0.019850260799741417
Ecriture standardisée	$T = 0.94 \pm 0.02 \text{ s}$	

d. Expériences sans variabilité observée (incertitude de type B)

Il arrive parfois que la variabilité des mesures ne soit pas accessible :

- soit parce qu'on ne dispose que d'une observation unique (par manque de temps ou parce qu'il n'est pas matériellement possible de reproduire le processus de mesure) ;
- soit parce que la répétition des observations conduit systématiquement à la même valeur, la variabilité intrinsèque étant masquée par la résolution limitée de l'appareil employé.

Exemple : la masse que nous avons suspendu ressort porte l'indication $m = 100 \text{ g}$.

La difficulté est alors d'estimer la variabilité intrinsèque d'une mesure sans possibilité de l'observer. Même si nous disposions d'une balance précise au dixième de gramme près, il est probable qu'on lise toujours $m = 100,0 \text{ g}$.

Pour évaluer une telle incertitude dite " de type B ", il faut distinguer plusieurs cas.

Premier cas : la documentation de l'appareil de mesure n'est pas accessible et l'expérimentateur ne dispose d'aucune autre information qu'une limite inférieure et une limite supérieure pour la grandeur X étudiée, c'est-à- dire qu'il peut définir un intervalle $[x_{\min}; x_{\max}]$ au sein duquel il semble "raisonnable" de situer la valeur x recherchée.



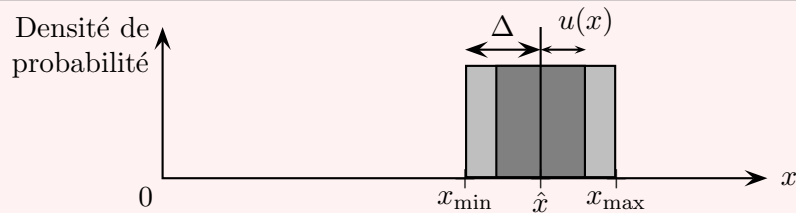
Exemple : incertitude de lecture sur une échelle graduée, incertitude de repérage ...

Incertitude de type B avec limites inférieure et supérieure identifiées

- Le meilleur estimateur \hat{x} est au centre de l'intervalle de largeur $2\Delta = x_{\max} - x_{\min}$

$$\hat{x} = \frac{x_{\min} + x_{\max}}{2} \quad \text{et} \quad \Delta = \frac{x_{\max} - x_{\min}}{2}$$

- On fait l'hypothèse d'une répartition équiprobable du résultat de la mesure en tout point de cet intervalle (distribution dite "uniforme" ou "rectangulaire").



Remarque : on peut établir que $u(x) = \frac{\Delta}{\sqrt{3}}$ mais cette relation n'est pas à retenir.

Méthode :




```
np.random.uniform(xmin,xmax,N)   ou   x+Delta*np.random.uniform(-1,+1,N)
```

génère un tableau de N valeurs réparties de façon uniforme entre min et max et d'écart type $\frac{\Delta}{\sqrt{3}}$.

Exemple : compléter le notebook pour m , la masse suspendue avec

```
np.random.uniform(mmin,mmax,N)   ou   m+Delta*np.random.uniform(-1,+1,N)
```

qui génère un tableau de N valeurs réparties uniformément entre m_{min} et m_{max} .

Visualisez la distribution et calculez son incertitude type. 

Deuxième cas : la documentation de l'appareil de mesure est disponible.

Incertitude de type B avec la notice de l'appareil de mesure :

- Si le constructeur de l'appareil de mesure fournit directement l'incertitude type, on se contente alors de récupérer cette valeur de $u(x)$ et on prend une distribution normale.
- Si le constructeur fournit une "incertitude constructeur" ou une "précision" Δ , sans autre explicitation alors on considère une loi de distribution uniforme.

4. Comparaison de deux résultats de mesure

Pour comparer deux résultats de mesure entre eux (issus par exemple de deux protocoles expérimentaux différents) il faut disposer d'un critère quantitatif permettant d'indiquer si ces deux mesurages sont compatibles ou non.

a. Ecart normalisé (z - score)

Définition

Lorsque deux processus de mesurage d'une même grandeur x conduisent à des résultats $\hat{x}_1 \pm u(x_1)$ et $\hat{x}_2 \pm u(x_2)$, on appelle écart normalisé la grandeur sans dimension suivante :

$$E_n = \frac{|\hat{x}_1 - \hat{x}_2|}{\sqrt{u^2(x_1) + u^2(x_2)}}$$

Critère quantitatif de comparaison

Par convention, on dira que deux processus de mesurage sont compatibles si leur écart normalisé E_n est inférieur à 2.

Remarque : l'écart normalisé est parfois également nommé "z - score".

b. Justification graphique et mise en application

On représente ci-dessous les histogrammes associés aux distributions des valeurs raisonnablement attribuables aux deux mesurages différents. Lorsque le recouvrement de ces deux distributions est suffisamment important (écart normalisé $E_n < 2$), les deux résultats sont jugés compatibles.

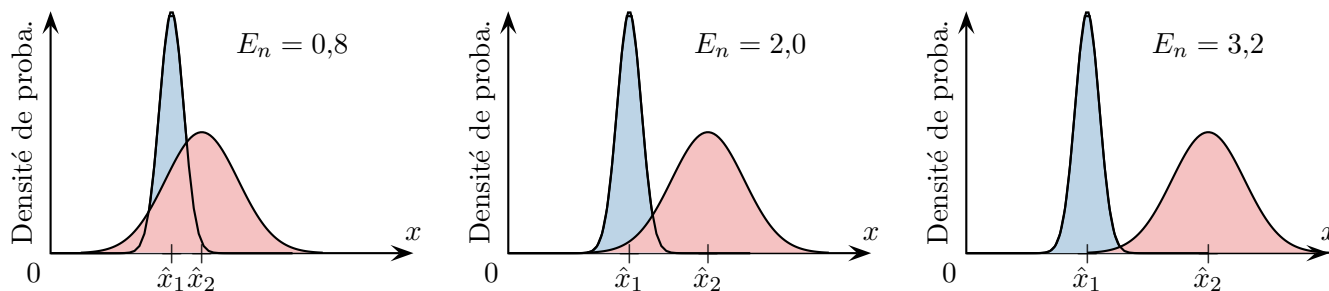


FIGURE 5 – Interprétation graphique de l'écart normalisé entre deux mesures

Exemple : comparez la période que vous avez mesurée à celle de votre groupe voisin 

Remarque : pour comparer une valeur mesurée à une "valeur de référence" (une constante connue par exemple) $x_{\text{réf}}$, on procède de la même manière en considérant une incertitude type nulle pour cette valeur "parfaitement connue" : $E_n = \frac{|\hat{x} - x_{\text{réf}}|}{u(x)}$

c. Que faire en cas d'incompatibilité ?

Il arrive couramment d'obtenir $E_n > 2$. Comment interpréter ?

- Il est possible que l'incertitude type ait été sous-estimée et il convient donc de réexaminer les choix qui ont mené à son évaluation.
- Il est possible que le processus de mesurage n'ait pas été correctement réalisé (mesure d'un rayon au lieu d'un diamètre, erreur de parallaxe lors de la lecture du volume sur une éprouvette graduée, erreurs d'arrondis dans les calculs, composant électronique défectueux, ...).
- La modélisation envisagée du phénomène observé est incomplète (loi physique valable dans un domaine de paramètres plus restreint que celui qu'on explore en pratique).

Si on ne comprend pas d'où vient un désaccord, ou si on n'est pas certain de l'origine de ce désaccord, il peut être judicieux de reprendre le mesurage.

Dans tous les cas, la présence d'une incompatibilité n'est pas forcément synonyme d'échec. La méthode de mesure peut tout à fait légitimement donner des résultats incompatibles.

En fin de compte, **l'important est le soin accordé au mesurage, ainsi que l'honnêteté des observations décrites par l'expérimentateur.**

Remarque : ici par exemple il est tout à fait possible que les résultats ne soient pas compatibles puisqu'il ne s'agit pas du même ressort.

II Composition des incertitudes

1. Contexte et mise en application

Très souvent, la grandeur physique qui nous intéresse n'est pas directement accessible par mesurage mais peut s'obtenir en combinant entre elles d'autres grandeurs physiques mesurables.

Exemple : Lors de l'étude théorique du système masse ressort vertical, on peut montrer que la période d'oscillation est reliée à la masse m suspendue et à k , une constante caractéristique du ressort appelée constante de raideur, par la relation

$$\frac{2\pi}{T} = \sqrt{\frac{k}{m}} \Rightarrow k = \frac{4\pi^2 m}{T^2}$$

On cherche le lien entre l'incertitude type $u(k)$ de la grandeur calculée et les incertitude types $u(T)$ et $u(m)$ évaluées expérimentalement.

2. Approche mathématique, loi de composition des incertitudes

Les valeurs expérimentales de différentes grandeurs physiques sont supposées parfaitement connues (mesurées au préalable) : $X_1 = \hat{x}_1 \pm u(x_1)$; $X_2 = \hat{x}_2 \pm u(x_2)$; ...

Que peut-on alors en déduire pour une grandeur calculée du type $Y = f(X_1, X_2, \dots)$ où f est une fonction connue ?

L'évaluation du meilleur estimateur de cette grandeur Y est immédiate : $\hat{y} = f(\hat{x}_1, \hat{x}_2, \dots)$ mais quelle est l'incertitude type correspondante $u(y)$? Les deux seuls cas à connaître sont les suivants.

”**Somme**” : pour une grandeur calculée de la forme $Y = \alpha_1 X_1 + \alpha_2 X_2 + \dots$ (où les α_i constantes),

$$u(y) = \sqrt{\alpha_1^2 u^2(x_1) + \alpha_2^2 u^2(x_2) + \dots}$$

Somme ou différence

$$Y = X_1 \pm X_2 \Rightarrow u(y) = \sqrt{u^2(x_1) + u^2(x_2)}$$

”**Produit**” : pour une grandeur calculée de la forme $Y = \lambda X_1^{\alpha_1} \cdot X_2^{\alpha_2} \dots$ (où λ et les α_i constantes),

$$\frac{u(y)}{|\hat{y}|} = \sqrt{\left(\alpha_1 \frac{u(x_1)}{\hat{x}_1}\right)^2 + \left(\alpha_2 \frac{u(x_2)}{\hat{x}_2}\right)^2 + \dots}$$

Produit ou quotient

$$Y = X_1 X_2^{\pm 1} \Rightarrow \frac{u(y)}{|\hat{y}|} = \sqrt{\left(\frac{u(x_1)}{\hat{x}_1}\right)^2 + \left(\frac{u(x_2)}{\hat{x}_2}\right)^2}$$

Remarque : il y a des restrictions dans l'utilisation des formules mathématiques précédentes !

- Les relations précédentes imposent que les grandeurs X_1 , X_2 , ... soient indépendantes entre elles (autrement dit que les variations de l'une n'influent pas sur les autres).
- Dans le cas d'une composition de type ”PRODUIT” (fonction f non-linéaire), la relation énoncée n'est valable que si les incertitudes type relatives $u(x_i)/\hat{x}_i$ restent modestes (c'est-à-dire de l'ordre du pourcent).

Exemple : dans le cas de l'oscillateur mécanique précédent on appliquerait la relation

$$k = \frac{4\pi^2 m}{T^2} = 4\pi^2 m T^{-2} \Rightarrow u(k) = k \sqrt{\left(\frac{u(m)}{m}\right)^2 + \left(2 \frac{u(T)}{T}\right)^2} = \frac{4\pi^2 m}{T^2} \sqrt{\left(\frac{u(m)}{m}\right)^2 + \left(2 \frac{u(T)}{T}\right)^2}$$

L'application numérique donne $u(k) \simeq 0,50 \text{ kg}\cdot\text{s}^{-2}$ et on écrira

$$k = 8,53 \pm 0,50 \text{ kg}\cdot\text{s}^{-2}$$

3. Recours à une simulation Monte Carlo

Au lieu de travailler avec des calculs lourds on utilise la puissance de Python.

Méthode :

Dans le cas d'une relation de la forme $Y = f(X_1, X_2, \dots)$, la mise en œuvre de la simulation Monte Carlo repose alors sur les étapes suivantes.

- ① On fixe un nombre N de tirages à réaliser (avec $N \gg 1$ afin de traduire le plus fidèlement possible les variabilités des grandeurs X_1, X_2, \dots).

- ② On effectue un tirage aléatoire pour obtenir les valeurs possibles de $X_1, X_2 \dots$
- ③ On calcule les valeurs correspondantes de la grandeur Y .
- ④ La meilleure estimation \hat{y} de la valeur de la grandeur Y correspond alors à la moyenne de cet ensemble et l'incertitude type $u(y)$ à son écart type.

Exemple : déterminer la valeur k de la constante de raideur de votre ressort et comparer à celle donnée par le constructeur ou à celle des groupes voisins 

III Validation d'une loi affine

1. Contexte et mise en application

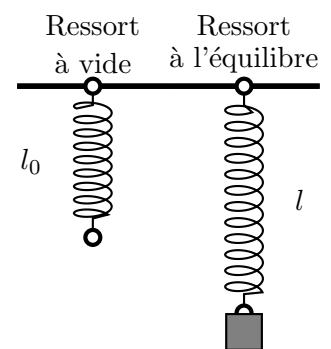
Il est assez fréquent en physique et en chimie qu'un modèle théorique prédise une relation affine entre des grandeurs physiques mesurables X et Y , de la forme $Y = aX + b$.

Exemple : on souhaite vérifier la loi de Hooke qui stipule que la déformation des solides élastiques est une fonction linéaire des contraintes.

Dans le cas d'un ressort vertical soumis au poids d'une masse m à l'équilibre, l'application de la première loi de Newton implique

$$l = \frac{mg}{k} + l_0$$

où l_0 est la longueur "à vide" et k la constante de raideur du ressort.



2. Régression linéaire

Méthode :




Pour tester l'adéquation d'un modèle théorique $Y = aX + b$ avec l'expérience, on réalise n mesures de couples de valeurs (x_i, y_i) puis on utilise une "régression linéaire" qui consiste à trouver les coefficients a et b s'accordant "au mieux" avec une relation du type $Y = aX + b$.

On utilise la commande

```
a,b = np.polyfit(x,y,1)
```

appliquée aux tableaux x et y

Remarque : en toute rigueur, il faudrait plutôt parler de "régression affine".

Exemple : réalisez la prise de mesure (m, l) pour une dizaine de masses différentes et complétez le notebook pour déterminer les coefficients a et b . 

3. Un premier contrôle "visuel"

Méthode :

- ① On représente graphiquement les n points de mesure $\bullet\{(x_i, y_i)\}$ à l'aide de la commande

```
plt.plot(x,y,'bo',label='Mesures')
```

- ② On fait apparaître sur le graphe la droite de régression optimale $a.x + b$.

```
plt.plot(x,a*x+b,'r',label='Droite de régression optimale')
```

- ③ Les données expérimentales sont en accord avec le modèle si les points de mesure apparaissent bien alignés et que la droite de régression passe à proximité de ces points.

Exemple : placez les points de mesure (m,l) sur le notebook ainsi que la droite optimale de régression linéaire. Le modèle vous semble-t-il (à priori) validé?

4. Recours à une simulation Monte Carlo

Si on souhaite en plus déterminer les incertitudes type $u(a)$ et $u(b)$ on peut utiliser la méthode de simulation Monte Carlo ci-dessous.

Méthode :

- ① On fixe un nombre N de simulations à réaliser (avec $N \gg 1$).
- ② A chaque itération (pour k allant 1 à N) :
- on effectue un tirage aléatoire des n couples de valeurs (x_i, y_i) , en adoptant les lois de distribution adéquates pour X et Y ;
 - pour cet ensemble de données $\{(x_1, y_1), \dots, (x_n, y_n)\}$, on effectue une régression linéaire conduisant à identifier la pente a_k optimale et l'ordonnée à l'origine b_k optimale.
- ③ On effectue ensuite une étude statistique de ces ensembles $\{a_k\}$ et $\{b_k\}$ pour déterminer le meilleur estimateur (moyenne arithmétique) et l'incertitude type (écart type).

$$a = \bar{a} \pm u(a) \text{ et } b = \bar{b} \pm u(b)$$

Exemple : appliquer cette méthode dans le notebook. Remarquez en rouge l'ensemble des régressions linéaires potentielles.

5. Validation d'une régression linéaire

On ajoute les incertitudes type et on trace la droite optimale $y = \hat{a}.x + \hat{b}$ sur le graphique.

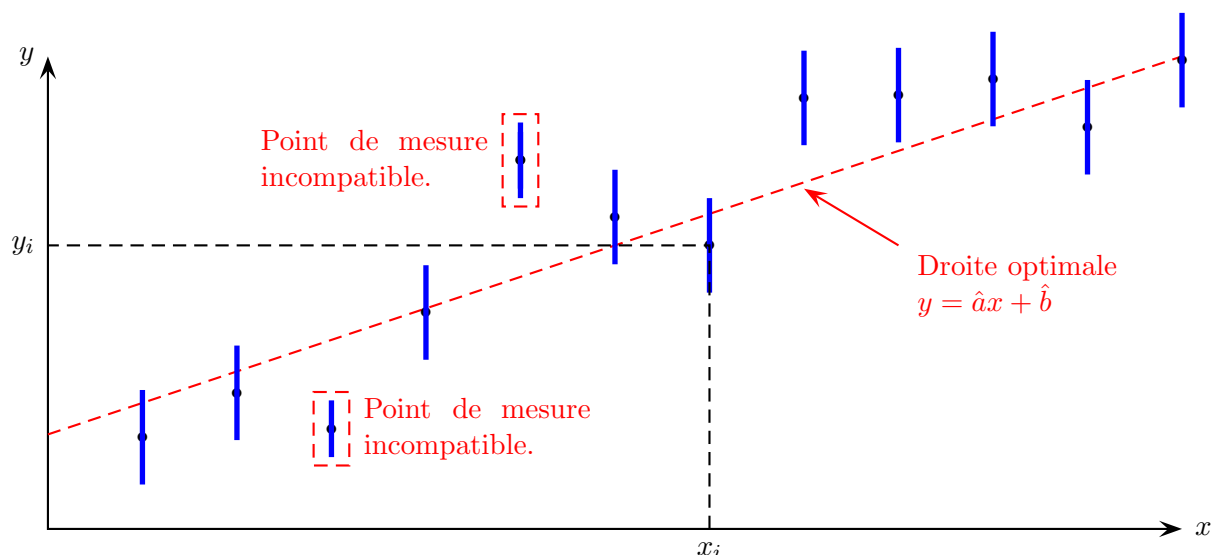



FIGURE 6 – Résultat d’une régression linéaire

On note "résidu" la distance algébrique $\delta_i = y_i - (\hat{a}x_i + \hat{b})$ entre les données expérimentales et la droite modèle. On a alors $E_{n,i} = \frac{|\delta_i|}{u(y_i)}$ l'écart normalisé pour chaque point de figure.

Critère de validation

L'ajustement affine des données expérimentales est jugé cohérent si $E_{n,i} < 2$ c'est à dire si tous les résidus restent inférieurs au double de l'incertitude type associée.


Exemple : compléter le notebook en utilisant la commande `plt.errorbar()`.

Si le modèle est validé, déterminez à nouveau la constante de raideur k du ressort et comparez à la valeur obtenue par la méthode précédente. Conclure. 

6. Cas d'un modèle vraiment linéaire

Supposons à présent qu'un modèle théorique prédise une relation effectivement linéaire entre des grandeurs mesurables X et Y , c'est-à-dire de la forme $Y = aX$.

Il pourrait être tentant de procéder comme précédemment mais la fonction `np.polyfit()`, conçue pour des relations affines donne parfois des résultats aberrants.

Exemple : en nommant $d = l - l_0$ la déformation du ressort précédent, la loi de Hooke s'écrit maintenant $d = \frac{g}{k}m$. Tracez $d(m)$ dans le Notebook et essayez d'en déduire k par la méthode précédente. 

Méthode :



On peut revenir à une **approche statistique** de type A :

- ① On calcule le paramètre $a_i = \frac{y_i}{x_i}$ pour chacun des n points de mesure.
- ② On effectue alors une étude statistique de cet ensemble $\{a_i\}$ pour déterminer la meilleure estimation (moyenne arithmétique) et l'incertitude type (écart type) :

$$a = \bar{a} \pm u(a)$$


Exemple : utilisez cette méthode dans le Notebook pour calculer à nouveau k et comparer à la valeur obtenue par la première méthode. 

Table des matières

I Valeur mesurée et incertitude type (mesurage)

1. Variabilité de la mesure d'une grandeur physique
2. Meilleur estimateur et incertitude type associés à un mesurage
3. Approche algorithmique, méthode de Monte Carlo
 - a. Principe
 - b. Expériences avec variabilité observée (incertitude de type A)
 - c. Ecriture du résultat d'une mesure
 - d. Expériences sans variabilité observée (incertitude de type B)
4. Comparaison de deux résultats de mesure
 - a. Ecart normalisé (z - score)
 - b. Justification graphique et mise en application
 - c. Que faire en cas d'incompatibilité ?

II Composition des incertitudes

1. Contexte et mise en application
2. Approche mathématique, loi de composition des incertitudes
3. Recours à une simulation Monte Carlo

III Validation d'une loi affine

1. Contexte et mise en application
2. Régression linéaire
3. Un premier contrôle " visuel "
4. Recours à une simulation Monte Carlo
5. Validation d'une régression linéaire
6. Cas d'un modèle vraiment linéaire

Les savoirs :

- ♥ Différence entre incertitude de type A et de type B.
- ♥ Différence entre loi normale et uniforme.
- ♥ Notion d'écart normalisé : formule et interprétation.
- ♥ Formule de composition des incertitudes dans les cas simples (somme ou différence / produit ou quotient).

Les savoir-faire :

- ✓ Identifier et évaluer l'incertitude type sur une grandeur mesurée.
- ✓ Simuler à l'aide de python un processus aléatoire pour calculer l'écart type d'une grandeur physique ou pour vérifier un modèle (regression linéaire).