

I – Mesures et incertitudes

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.

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 contenus	Capacités exigibles	Remarques
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. <i>Capacité numérique : 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.</i>	Après avoir reconnu l'ensemble des causes d'incertitudes. On simplifie le calcul en ne gardant que les prépondérantes.
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.	On nomme l'écart normalisé z-score.
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. <i>Capacité numérique : 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>	

I – Variabilité de la mesure

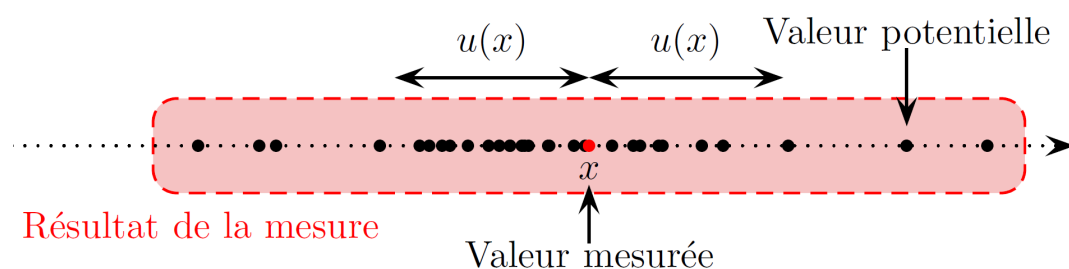
I-1) Incertitude-type

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 à une dispersion des diverses valeurs observées. Cette variabilité est naturelle et fait intrinsèquement 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. L'incertitude constitue alors une indication de la dispersion de cet ensemble de valeurs.

Incertitude-type : la quantification de la variabilité potentielle de la mesure x d'une grandeur physique X est appelée incertitude-type et est notée $u(x)$.

Par définition, l'incertitude-type correspond à l'écart-type de la distribution des données qui seraient issues d'un grand nombre de répétitions de la mesure.



L'incertitude-type caractérise la variabilité de l'ensemble du processus de mesurage : nature du protocole expérimental mis en œuvre, matériel utilisé, conditions expérimentales et l'habileté de l'expérimentateur.

I-2) Les sources de variabilité

a) Les différentes sources

Les incertitudes peuvent être liées à l'opérateur, l'environnement, aux instruments ou à la méthode de mesure.

Soit x une grandeur mesurée et u_x ou $u(x)$ l'incertitude type sur x liée à la mesure, alors u_x résulte de la composition des incertitudes-type suivantes :

- u_c : incertitude type constructeur liée à la qualité de l'appareil.
- u_{lec} : incertitude type de lecture (ou de repérage) lié à la lecture sur l'instrument (graduations, ...).
- $u_{spé}$: incertitude type spéciale liée à la difficulté de faire la mesure : pointés, encadrement, non linéarité...

Par exemple, lorsque l'on mesure sur un banc d'optique la distance entre la position d'une image (à travers une lentille) formée sur un écran, on a une première source d'incertitude qui vient de la lecture de la position de l'écran sur le banc gradué. On lui associe l'incertitude-type $u_1(x)$. (Incertitude de lecture).

Une deuxième source d'incertitude vient du fait qu'il n'existe pas qu'une seule position de l'écran on l'on peut dire que l'image est nette, mais qu'au contraire tout un intervalle de position de l'écran est acceptable. On associe l'incertitude-type $u_2(x)$ à cette deuxième source d'incertitude sur la position de l'image. (Incertitude de pointé).

Une troisième source d'incertitude est liée à la qualité des graduations sur le banc d'optique $u_3(x)$. (Incertitude constructeur)

Il faut ensuite comparer qualitativement les ordres de grandeur des différentes incertitudes-types identifiées et essayer de n'en garder qu'une seule (considérant donc que les autres sont négligeables). Si ce n'est pas possible, i.e. que les différentes incertitudes-types sont toutes du même ordre de grandeur, alors l'incertitude-type globale sera obtenue par une somme quadratique des différentes contributions : $u_x^2 = u_c^2 + u_{lec}^2 + u_{spé}^2$

I-3) Valeur mesurée et incertitude type

La valeur mesurée x est la meilleure estimation possible de la grandeur X mesurée. Comment la déterminer ?

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

On demande à un élève de prélever 100,0 mL d'eau à l'aide d'une fiole jaugée (préalablement tarée) et on mesure la masse d'eau pour chaque essai. On obtient les résultats suivants :

	Essai n° 1	Essai n° 2	Essai n° 3	Essai n° 4	Essai n° 5
Masse relevée (g)	99,6	99,5	99,9	99,8	99,4

La valeur attribuée à la masse d'eau est :

$$m = \frac{99,6 + 99,5 + 99,9 + 99,8 + 99,4}{5} = 99,6 \text{ g.}$$

Si on ne dispose que d'une valeur unique, alors c'est celle que l'on prend.

Se pose alors la question de la détermination de l'incertitude-type attachée à cette valeur.

II – Évaluation d'une incertitude-type par approche statistique (type A)

II-1) Principe

Il s'agit de répéter N fois une expérience dans les mêmes conditions (même expérimentateur, même matériel, même protocole) : l'évaluation de l'incertitude-type sur la valeur de la grandeur mesurée est effectuée à l'aide d'un traitement statistique des N observations.

Le meilleur estimateur de la valeur de la grandeur physique X correspond à la moyenne arithmétique des valeurs de x (N est le nombre de mesures répétées) $\bar{x}$ défini par :

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{x_1 + x_2 + \dots + x_N}{N}$$

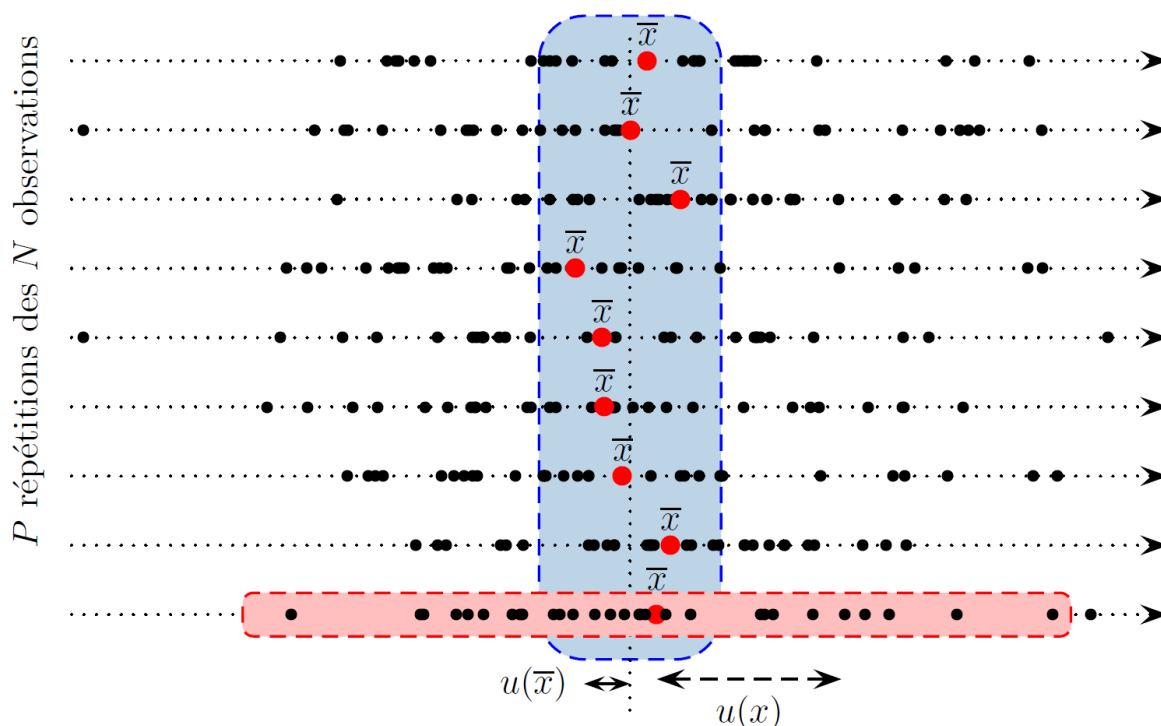
Il est important de noter ici le changement de point de vue : l'objectif n'est plus simplement d'évaluer l'incertitude-type $u(x)$ associée à un mesurage unique, mais plutôt d'évaluer l'incertitude-type $u(\bar{x})$ sur la mesure de la moyenne $\bar{x}$ déduite de N observations. Cette expérience est différente d'une observation unique et possède donc une incertitude-type propre :

$$u(\bar{x}) < u(x)$$

II-2) Incertitude-type A

Comment estimer la variabilité potentielle de cette mesure de la moyenne $\bar{x}$? Une première solution pour évaluer l'incertitude-type $u(\bar{x})$ consiste à procéder comme précédemment :

- On recommence P fois la réalisation de N observations de façon à obtenir autant d'évaluations de la moyenne arithmétique $\bar{x}$.
- Par définition, l'incertitude-type $u(\bar{x})$ correspond alors à l'écart-type de cet ensemble de valeurs.



Cette méthode étant chronophage, on s'appuiera sur un théorème mathématique qui stipule que l'incertitude-type de la moyenne est liée à l'écart-type expérimental d'une série de mesure par :

$$u(\bar{x}) = \frac{\sigma(x)}{\sqrt{N}}$$

où l'écart type expérimental $\sigma(x)$ déduite des N mesures $x_1, x_2, \dots, x_N$ est donnée par :

$$\sigma(x) = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Incertain type A :

Lorsqu'on réalise N observations de la valeur d'une grandeur physique X avec variabilité observée, on considère que :

- Le meilleur estimateur de la grandeur X correspond à la moyenne $\bar{x}$ de cet ensemble de valeurs observées.
- L'incertitude-type de cette moyenne vérifie :

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

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

III – Évaluation d'une incertitude-type pour une mesure unique (type B)

III-1) Principe

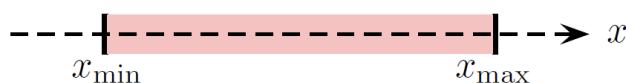
On suppose n'avoir réalisé qu'une seule observation (ou plusieurs qui aboutissent à la même valeur) : il s'agit alors d'une incertitude de type B. Dans ce cas, la variabilité de la mesure est estimée par un modèle probabiliste (« quelle probabilité la mesure a-t-elle de donner telle valeur ? ») :

- Dans des cas très rares, le constructeur de l'appareil de mesure utilisé fournit directement l'incertitude-type associée à la valeur mesurée ;
- Parfois, le fabricant indique une « précision » ou une « tolérance » (termes vagues et mal définis) qui permettent d'établir un intervalle de valeurs constituant le résultat de la mesure.

Mais la plupart du temps, la source d'incertitude prédominante est celle du repérage (pointage) effectué par l'opérateur : il est donc nécessaire d'estimer un intervalle de valeurs à l'intérieur duquel on est sûr que la grandeur mesurée appartient.

III-2) Intervalle de mesure

Considérons la situation expérimentale suivante : 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.



Cette situation peut correspondre aux cas usuels suivants :

- Incertitude de lecture : la valeur mesurée est lue sur une échelle graduée. La largeur $x_{max} - x_{min}$ de l'intervalle correspond alors à une

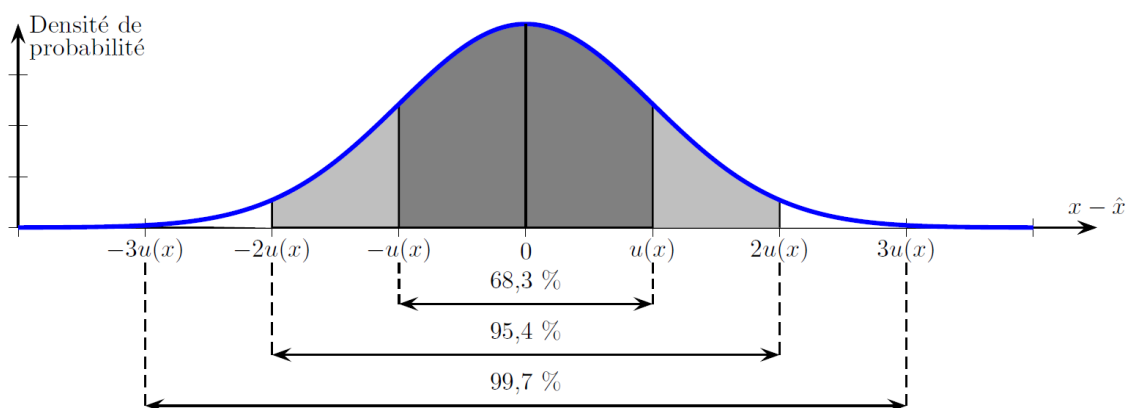
graduation, voire à une demi-graduation.

- Incertitude de repérage : toute une plage de valeurs semble convenir :
 - Latitude de mise au point avec une lentille,
 - Repérage d'une résonance en électrocinétique,
 - Fluctuations temporelles de la valeur affichée par l'appareil de mesure...

III-3) Loi de probabilité normale

Si l'on devait effectuer un très grand nombre de fois la mesure de la grandeur, on observerait naturellement une distribution des valeurs mesurées suivant une « courbe en cloche ». La loi de probabilité correspondante est la loi normale (ou loi gaussienne).

Cette distribution est caractérisée par sa valeur moyenne et son écart-type $u(x)$.



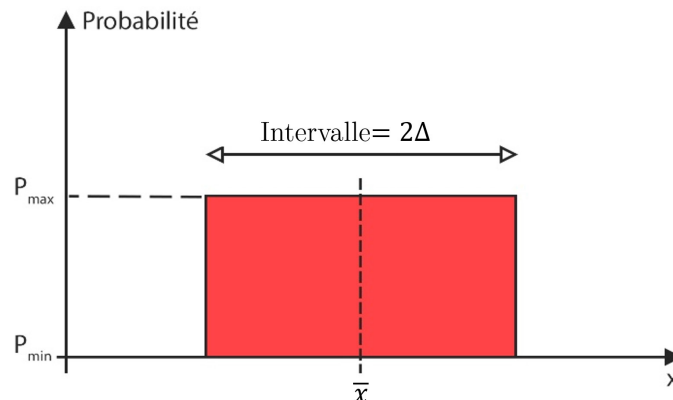
On montre que si l'on effectue une seule mesure :

- La probabilité de trouver la valeur dans l'intervalle $[\bar{x} - u(x); \bar{x} + u(x)]$ vaut 68%
- La probabilité de trouver la valeur dans l'intervalle $[\bar{x} - 2u(x); \bar{x} + 2u(x)]$ vaut 95%
- La probabilité de trouver la valeur dans l'intervalle $[\bar{x} - 3u(x); \bar{x} + 3u(x)]$ vaut 99%

Ce pourcentage est appelé niveau de confiance.

III-4) Loi de probabilité rectangulaire (uniforme)

Pour calculer l'écart-type de l'ensemble des valeurs comprises dans l'intervalle défini, on fait le plus souvent l'hypothèse que ces valeurs y sont équiréparties (on parle de distribution uniforme). Autrement dit, si on représentait un histogramme avec un grand nombre de valeurs dans cet intervalle, celui-ci serait rectangulaire. Ce n'est pas la distribution la plus naturelle des valeurs possibles, mais c'est le modèle probabiliste le plus simple.



A l'aide d'un calcul de variance sur la distribution on peut montrer que l'incertitude-type liée à la distribution « uniforme » s'écrit :

$$u(x) = \frac{\text{intervalle}}{\sqrt{12}} = \frac{\Delta}{\sqrt{3}}$$

III-5) Incertitude-type B

Incrtitude type B :

Lorsqu'on identifie un intervalle $[x_{min}; x_{max}]$ au sein duquel il semble raisonnable de situer la valeur x d'une grandeur physique X , le meilleur estimateur correspond au centre de l'intervalle :

$$\bar{x} = \frac{x_{min} + x_{max}}{2}$$

L'intervalle peut alors s'écrire $[\bar{x} - \Delta; \bar{x} + \Delta]$ en introduisant la demi-largeur : $\Delta = \frac{x_{max} - x_{min}}{2}$

Dans le cas d'une distribution uniforme, on peut établir que l'incertitude-type associée au mesurage vaut :

$$u(x) = \frac{\Delta}{\sqrt{3}}$$

III-6) Présence d'une notice constructeur

Lorsque la notice de l'appareil de mesure est disponible, plusieurs cas de figure sont possibles.

Le constructeur de l'appareil de mesure fournit directement l'incertitude-type : on se contente alors de récupérer cette valeur de $u(x)$.

Le constructeur fournit une incertitude constructeur ou une précision, sans davantage d'explicitation sur la signification de ce paramètre. Dans ce cas on considère que la précision évoquée s'identifie à la demi-étendue d'un intervalle de valeurs possibles $[\bar{x} - \Delta; \bar{x} + \Delta]$ centré sur la valeur lue x et associé à une loi de distribution uniforme. L'incertitude-type vaut alors :

$$u(x) = \frac{\Delta}{\sqrt{3}}$$

Exemple :

Les appareils qu'on utilise au lycée sont très proches de ceux employés dans le mode professionnel ainsi le calcul de l'incertitude constructeur est très détaillé. Par exemple, la notice du Fluke 287 fournit ci-dessous propose une estimation de l'incertitude à l'aide de deux coefficients (les RLC mètre souvent à trois) a et b :

$$\Delta = a\% \times \text{Valeur}_{lue} + b \times U.R. \Rightarrow u(x) = \frac{a\% \times \text{Valeur}_{lue} + b \times U.R.}{\sqrt{3}}$$

Caractéristiques en tensions alternatives

Fonction	Plage	Résolution	Précision				
			20 à 45 Hz	45 à 65 Hz	65 Hz à 10 kHz	10 à 20 kHz	20 à 100 kHz]
AC mV	50 mV ^[1]	0,001 mV	1,5 % + 60	0,3 % + 25	0,4 % + 25	0,7 % + 40	3,5 % + 40 ^[5]
	500 mV	0,01 mV	1,5 % + 60	0,3 % + 25	0,4 % + 25	0,7 % + 40	3,5 % + 40
AC V	5 V ^[1]	0,0001 V	1,5 % + 60	0,3 % + 25	0,6 % + 25	1,5 % + 40	3,5 % + 40 ^[5]
	50 V ^[1]	0,001 V	1,5 % + 60	0,3 % + 25	0,4 % + 25	0,7 % + 40	3,5 % + 40
	500 V ^[1]	0,01 V	1,5 % + 60	0,3 % + 25	0,4 % + 25	Non spéc.	Non spéc.
	1 000 V	0,1 V	1,5 % + 60	0,3 % + 25	0,4 % + 25	Non spéc.	Non spéc.
dBV	-70 à -62 dB ^[3]	0,01 dB	3 dB	1,5 dB	2 dB	2 dB	3 dB
	-62 à -52 dB ^[3]	0,01 dB	1,5 dB	1,0 dB	1 dB	1 dB	2 dB
	-52 à -6 dB ^[3]	0,01 dB	0,2 dB	0,1 dB	0,1 dB	0,2 dB	0,8 dB
	-6 à +34 dB ^[3]	0,01 dB	0,2 dB	0,1 dB	0,1 dB	0,2 dB	0,8 dB
	34 à 60 dB ^[3]	0,01 dB	0,2 dB	0,1 dB	0,1 dB	Non spéc.	Non spéc.
Filtre passe-bas ^[4]			2 % + 80	2 % + 40	2 % + 10 -6 % -60 ^[2]	Non spéc.	Non spéc.
LoZ ^[4] V	1 000 V	0,1 V	2 % + 80	2 % + 40	2 % + 40 ^[6]	Non spéc.	Non spéc.

Si on mesure 1,2345V à 1kHz alors :

$$\Delta = \frac{0,6}{100} \times 1,2345 + 25 \times 0,0001 = 0,0099 \text{ V} \Rightarrow u(x) = 0,0057 \text{ V.}$$

IV – Présentation du résultat

V-1) Incertitude-type composée

On cherche souvent l'incertitude-type $u(x)$ sur la valeur d'une grandeur X calculée à partir d'une ou plusieurs grandeurs mesurées.

Supposons avoir évalué l'incertitude de plusieurs grandeurs physiques $(x_1, x_2, \dots, x_n)$ et que l'on souhaite déterminer la valeur d'une autre grandeur $y = f(x_1, x_2, \dots, x_n)$, qui dépend des valeurs de ces grandeurs, et évaluer l'incertitude $u(y)$ sur sa détermination. Les erreurs de mesure commises sur les x_i vont se répercuter sur y . Pour évaluer la nouvelle incertitude, on utilise pour cela la formule de propagation des incertitudes :

$$u(y) = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} u_{x_i} \right)^2}$$

Cette formule est simple à utiliser dans la plupart des cas que l'on rencontre cependant elle a des formes très simplifiées dans différentes opérations.

IV-2) Les différentes opérations

Type	Relation	Incertitude-type composée
Somme	$y = ax$	$u_y = a u_x$
	$y = ax \pm bz$	$u_y = \sqrt{a^2 u_x^2 + b^2 u_z^2}$
Produit	$y = a x^n$	$\frac{u_y}{ y } = n \frac{u_x}{ x }$
	$y = \frac{x}{z}$	$\frac{u_y}{ y } = \sqrt{\left(\frac{u_x}{x}\right)^2 + \left(\frac{u_z}{z}\right)^2}$
	$y = a \cdot w^\alpha \cdot x^\beta$	$\frac{u_y}{ y } = \sqrt{\alpha^2 \left(\frac{u_w}{w}\right)^2 + \beta^2 \left(\frac{u_x}{x}\right)^2}$

Lorsqu'on sort d'opérations classiques on pourra utiliser un algorithme utilisant la variabilité d'une mesure pour simuler un calcul d'incertitude.

IV-3) Recours à une simulation de Monte-Carlo

Pour mettre en œuvre une simulation Monte-Carlo, il faut connaître la loi de probabilité décrivant la variabilité des paramètres $X = \bar{x} \pm u(x)$ apparaissant dans l'expression de la grandeur calculée Y . On pourra généralement se restreindre aux cas suivants.

- Si le résultat de la mesure est issu d'une évaluation de type A, la loi de distribution peut être supposée normale (gaussienne centrée sur x_i et d'écart-type $u(x_i)$).
- Lorsque le résultat de la mesure est issu d'une évaluation de type B :
 - o Si la variabilité de la grandeur est caractérisée par un intervalle du type $[\bar{x}_i - \Delta_i; \bar{x}_i + \Delta_i]$, la loi de distribution est supposée uniforme (centrée sur $\bar{x}_i$, de demi-largeur Δ_i d'où $u(x_i) = \frac{\Delta_i}{\sqrt{3}}$)
 - o Si l'incertitude-type est directement indiquée dans la notice de l'appareil mais que la distribution sous-jacente n'est pas connue, on supposera que la distribution est normale (gaussienne centrée sur $\bar{x}_i$ et d'écart-type $u(x_i)$).

Méthode :

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

- i. On fixe un nombre N de simulations à réaliser (avec $N \gg 1$) afin de traduire le plus fidèlement possible les variabilités des grandeurs X_1 et X_2).
- ii. A chaque itération (pour k allant 1 à N) :
 - On effectue un tirage aléatoire de valeurs numériques (x_1, x_2) en adoptant les lois de distribution adéquates pour X_1 et X_2 .
 - On calcule la valeur correspondante de la grandeur Y : $y_k = f(x_1, x_2)$.
- iii. La meilleure estimation $\bar{y}$ de la valeur de la grandeur Y correspond alors à la moyenne de cet ensemble des $\{y_k\}$ et l'incertitude-type $u(y)$ à l'écart-type de cet ensemble.

IV-4) Présentation du résultat

Le résultat de la mesure d'une grandeur physique X doit inclure les informations suivantes :

- Le meilleur estimateur $\bar{x}$ de la valeur x de cette grandeur ;
- L'incertitude-type associée $u(x)$.

On peut noter ce résultat sous la forme suivante :

$$X = \bar{x} \pm u(x)$$

Il est important de préciser l'unité commune de ces deux paramètres. Il est également recommandé d'avoir recours à l'écriture scientifique lorsque cela paraît raisonnable.

L'incertitude-type résulte d'une estimation et possède donc forcément une valeur approchée. Dès lors :

- Donner un grand nombre de chiffres significatifs pour $u(x)$ risquerait de faire croire que l'incertitude-type est connue avec une grande précision;
- Donner un trop petit nombre de chiffres significatifs pourrait conduire à des erreurs d'arrondi dans les calculs menés ultérieurement.

On suggère donc d'arrondir systématiquement l'incertitude-type $u(x)$ à 2 chiffres significatifs.

Écriture des résultats :

- On arrondit l'incertitude-type $u(x)$ à deux chiffres significatifs.
- Le dernier chiffre de $\bar{x}$ possède la même position dans la mantisse (ou l'écriture décimale) que le dernier chiffre de l'incertitude-type.

Exemple :

Si on mesure 1,2345V à 1kHz avec un multimètre Fluke 287, on a vu que :
 $u(x) = 0,0057 \text{ V}$.

$$\begin{aligned} \Rightarrow U &= 1,2345 \pm 0,0057 \text{ V} \\ \text{ou} &= \underline{1234,5 \pm 5,7 \text{ mV}} \\ &\quad \text{5 et 7 : même décimale} \end{aligned}$$

V – Comparaison de deux mesures

V-1) Écart normalisé ou z-score

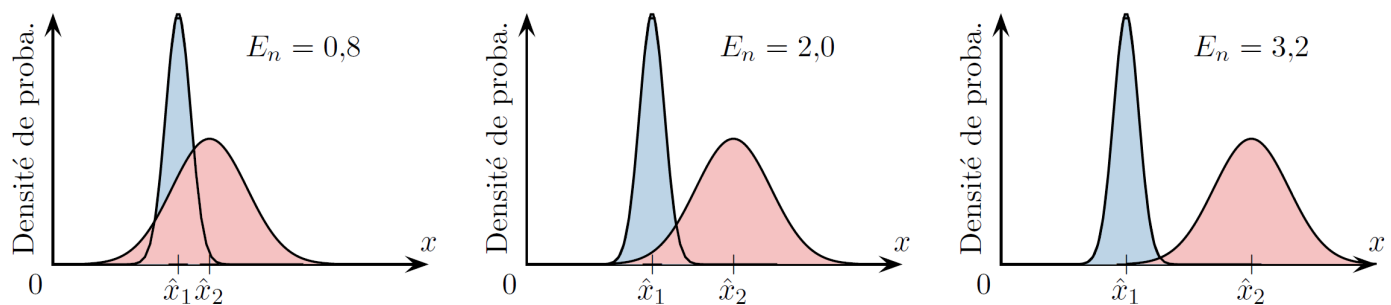
Pour pouvoir comparer deux mesures entre elles, il faut un critère quantitatif pour indiquer si ces deux mesures sont considérées comme compatibles ou incompatibles.

L'écart normalisé E_N ou z-score entre deux processus de mesure donnant les valeurs m_1 et m_2 et d'incertitudes types $u(m_1)$ et $u(m_2)$ est défini par :

$$E_N = \frac{|m_1 - m_2|}{\sqrt{u(m_1)^2 + u(m_2)^2}} = \frac{|m_1 - m_2|}{u(m_1 - m_2)}$$

Par convention, on qualifie souvent deux résultats de compatibles si leur écart normalisé vérifie la propriété $E_N \leq 2$.

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, les deux résultats sont jugés compatibles.



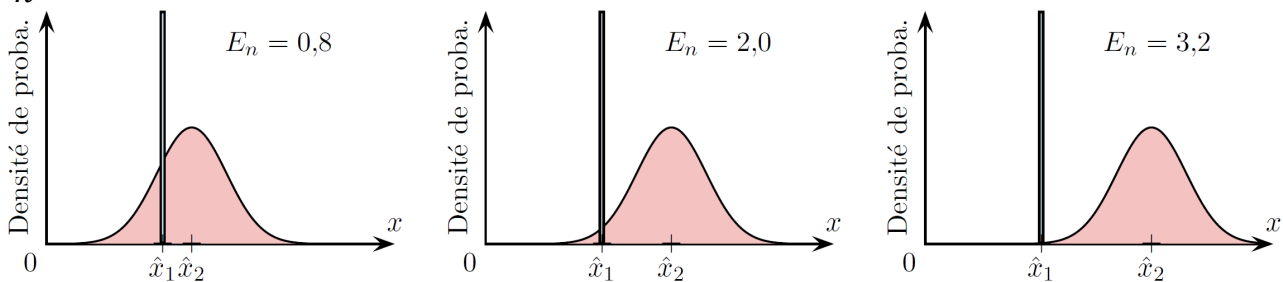
V-2) Comparaison d'un résultat de mesure à une valeur de référence

On appelle valeur de référence une valeur mesurée par une méthode dite de référence, c'est-à-dire par une méthode jugée scientifiquement supérieure à toute autre. Par extension, on pourra également appeler valeur de référence toute valeur mesurée dont l'incertitude-type peut être considérée négligeable devant celle obtenue par une autre méthode.

Pour comparer un résultat $\bar{x} \pm u(x)$ à une valeur de référence $x_{réf}$ (telle que son incertitude-type $u(x_{réf})$ soit négligeable devant $u(x)$), on considère l'écart normalisé :

$$E_N = \frac{|\bar{x} - x_{réf}|}{u(x)}$$

Le processus de mesurage est jugé compatible avec la valeur de référence $x_{réf}$ si $E_n \leq 2$.



On a représenté ci-dessus l'histogramme associé à la distribution des valeurs raisonnablement attribuables au mesurage. Lorsque le recouvrement entre cette distribution et la valeur de référence $x_{réf}$ est suffisamment important, les deux résultats sont jugés compatibles.

V-3) Incompatibilité des mesures

Comment interpréter une situation correspondant à $E_n > 2$?

- 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. (À l'inverse, il faut également éviter de surestimer l'incertitude-type car on risquerait alors de conclure par erreur à un accord entre des résultats de mesure pourtant incompatibles.
- 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, 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 ce désaccord, il peut être judicieux de reprendre le mesurage. Dans tous les cas, il faut bien comprendre que 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 et la complétude des observations décrites par l'expérimentateur.

VI – Régression linéaire

VI-1) Principe

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 . (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. Ce problème d'optimisation revient à minimiser la fonction suivante :

$$S(a, b) = \sum_{i=1}^N [y_i - (ax_i + b)]^2$$

La plupart des logiciels de traitement de données permettent d'identifier les valeurs optimales $\bar{a}$ et $\bar{b}$ de la pente et de l'ordonnée à l'origine pour un ensemble de données expérimentales $\{(x_i, y_i)\}$, mais sans prendre en compte les incertitudes expérimentales.

VI-2) Recours à une simulation de Monte-Carlo

Pour identifier les incertitudes-type $u(a)$ et $u(b)$ des paramètres d'une régression linéaire, on peut mettre en œuvre une simulation Monte-Carlo de la façon suivante.

Méthode :

- On fixe un nombre N de simulations à réaliser (avec $N \gg 1$).
- A chaque itération (pour k allant 1 à N) :
 - o 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 ;
 - o 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.
- 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 de façon à pouvoir écrire

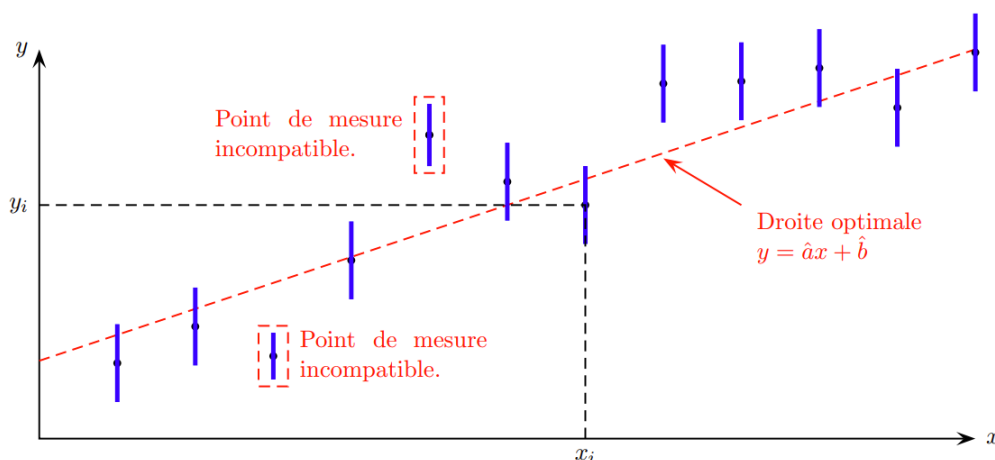
$$a = \bar{a} \pm u(a) \text{ et } b = \bar{b} \pm u(b)$$

VI-3) Validation de la régression linéaire

a) À l'aide des barres d'incertitude

La régression linéaire n'est pas une fin en soi-même, elle doit être suivie d'une discussion à l'issue de laquelle on validera (ou pas) le modèle affine proposé. On peut essayer un test visuel à l'aide des barres d'incertitude :

- On représente graphiquement les N points de mesure $\{(x_i, y_i)\}$, en faisant apparaître les barres d'incertitude correspondantes. On se placera dans le cas usuel où l'incertitude-type sur la grandeur X placée en abscisse est négligeable devant l'incertitude-type sur la grandeur Y placée en ordonnée (ce qui implique que $u(x_i) \ll u(y_i)$) pour chaque point de mesure). Sinon on devra tracer des « rectangles d'incertitude ».
- On fait apparaître sur ce graphe la droite de régression optimale (de paramètres $\bar{a}$ et $\bar{b}$).
- L'accord des données expérimentales avec le modèle est validé si les points de mesure apparaissent bien alignés et que la droite de régression passe à proximité suffisante de ces points. On rappelle que l'incertitude-type est une estimation de la variabilité de la mesure : il n'est donc pas anormal que des points expérimentaux soient éloignés de la valeur de la modélisation d'une ou deux incertitudes-type.



Si la représentation affine « traverse » toutes les barres d'incertitude, on validera le modèle. Cependant, afin d'avoir un aspect plus quantitatif, on va introduire les résidus.

b) À l'aide des écarts normalisés

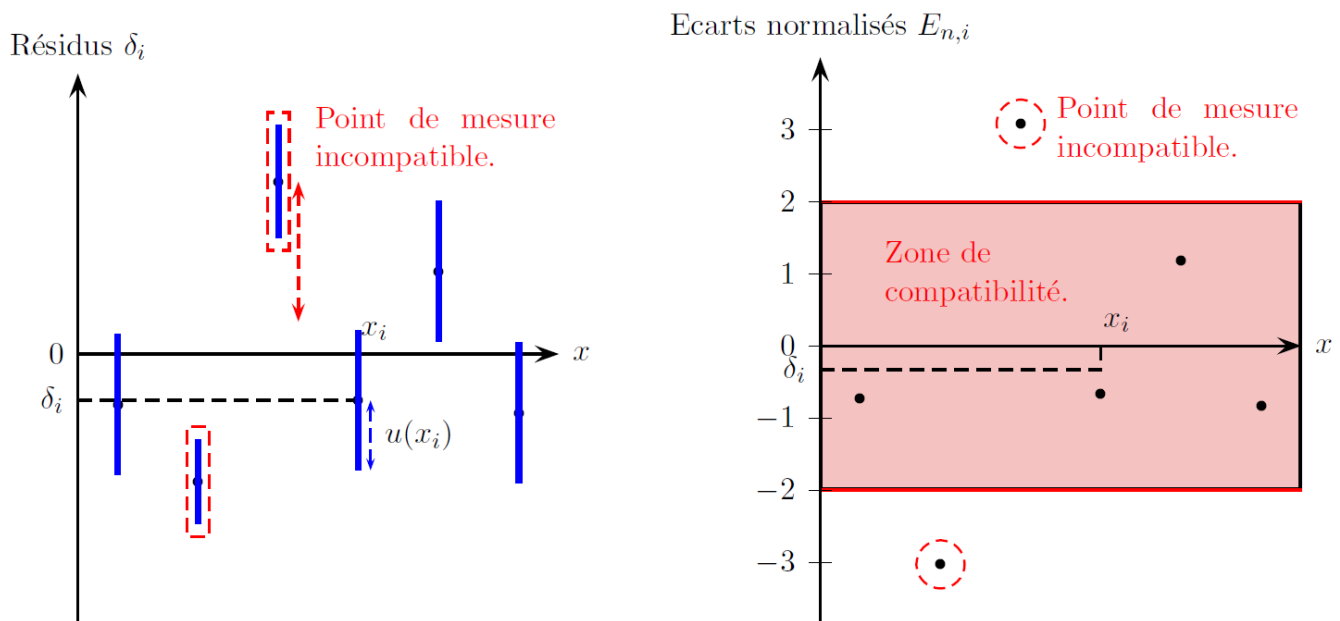
Pour faciliter l'analyse des résultats, on peut représenter les résidus, c'est-à-dire la distance algébrique $\delta_i = y_i - (\bar{a}x_i + \bar{b})$ entre les données expérimentales et la droite modèle. L'ajustement affine des données expérimentales est jugé cohérent si tous les résidus restent inférieurs au double de l'incertitude-type associée.

En rapportant les résidus à l'incertitude-type qui leur est associée, on peut également tracer les écarts normalisés :

$$E_{n,i} = \frac{\delta_i}{u(y_i)}$$

L'ajustement affine des données expérimentales est jugé cohérent si tous les écarts normalisés $E_{n,i}$ restent inférieurs à 2 en valeur absolue :

$$|E_{n,i}| \leq 2$$



Si l'éventualité d'un point aberrant n'est pas forcément à exclure, il faut à priori considérer que l'existence d'un seul point incompatible remet en cause la concordance l'ensemble des points de mesure avec une modélisation affine.

VII – Simulations numériques avec Python

VII-1) Incertitude-composé

Afin de mettre en évidence la simulation de Monte-Carlo sur les incertitudes composés, on va utiliser un exemple simple où l'incertitude composé est « difficile » à calculer avec les formules de base.

Dans le cadre d'une mesure de célérité d'une onde, les étudiants ont mesuré l'écart entre deux nœuds séparant dix longueurs d'ondes ainsi que la période. Ils ont déterminé les demi-étendues Δ attachées à chaque mesure.

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

# Entrez la période
T = 24.4e-6 # s
# Entrez la précision sur la période (demi-étendue de l'intervalle de confiance)
DeltaT = 0.1e-6 # s

# Positions de des points
A = 0.111 # m
B = 0.195 # m
# Entrez les précisions (demi-étendue de l'intervalle de confiance)
DeltaA = 0.005 # m
DeltaB = 0.005 # m

# Entrez la fonction de composition
def celerite(T,a,b):
    return (b-a)/(10*T)

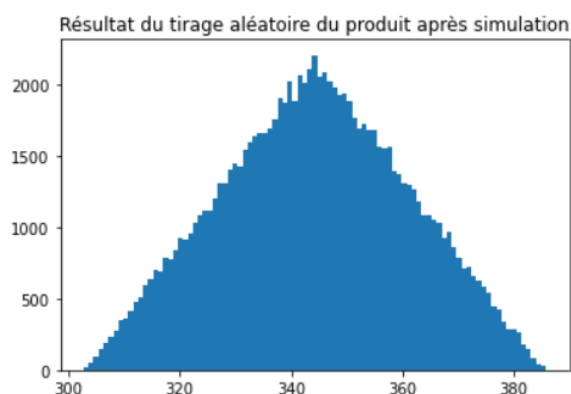
# Entrez le nombre de simulation que vous voulez effectuer
N = 100000

# Calculs avec une distribution de probabilité uniforme
Celerite=[]
for i in range(0,N):
    x = np.random.uniform(T-DeltaT,T+DeltaT)
    a = np.random.uniform(A-DeltaA,A+DeltaA)
    b = np.random.uniform(B-DeltaB,B+DeltaB)
    Celerite.append(celerite(x,a,b))

plt.figure(1)
plt.hist(Celerite,bins = 'rice')
plt.title('Résultat du tirage aléatoire du produit après simulation')
plt.xlabel("c (m/s)")

# Calcul et affichage moyenne et incertitude-type
moy = np.mean(Celerite)
std = np.std(Celerite,ddof=1)
print("Moyenne = {:.2f} m/s".format(moy))
print("Incertitude type = {:.2f} m/s".format(std))

Moyenne = 344.27 m/s
Incertitude type = 16.76 m/s
```



On retiendra l'utilisation des bibliothèques numpy et matplotlib afin d'accéder à des fonctions utiles pour l'obtention des grandeurs recherchés.

Si au lieu des demi-étendues, on aurait eu les incertitudes-type il aurait fallait multiplier par $np.sqrt(3)$ les DeltaX avant de faire appel à la distribution uniforme par $np.random.uniform$.

VII-2) Régression linéaire

a) Expérience

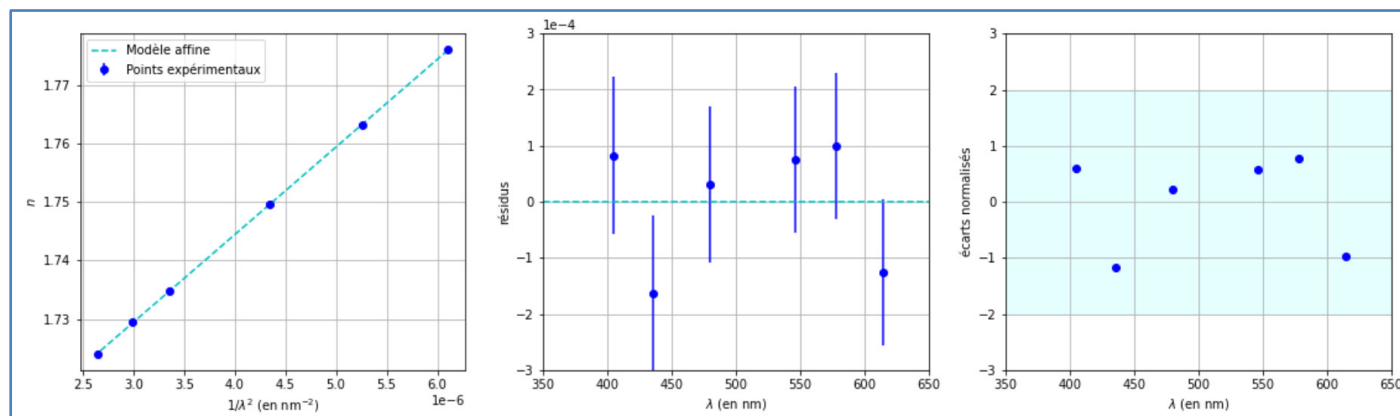
On dispose des valeurs de l'indice optique d'un prisme (en verre flint SF10) mesurées par spectroscopie pour différentes longueurs d'onde, ainsi que de leurs incertitudes-types associées.

λ (nm)	404,7	435,8	480,0	546,1	578,1	615,0
n	1,77610	1,76325	1,74957	1,73481	1,72942	1,72397
u(n)	0,00014	0,00014	0,00014	0,00013	0,00013	0,00013

On cherche à vérifier si le verre constituant le prisme suit la loi de Cauchy, c'est-à-dire si l'on peut écrire : $n = A + \frac{B}{\lambda^2}$ où A et B sont des coefficients positifs caractéristiques du verre.

b) Z-score

Le programme suivant permet de réaliser les étapes de la procédure de validation.



Grâce au troisième graphique, qui est le plus simple à interpréter, on peut en déduire que le modèle est validé par l'expérience.

```

# Importation des bibliothèques utiles

import numpy as np # pour le traitement vectoriel des données
import numpy.random as rd # pour les tirages aléatoires
import matplotlib.pyplot as plt # pour les graphes

# Saisie des valeurs expérimentales
lamb = np.array([404.7, 435.8, 480.0, 546.1, 578.1, 615]) # longueurs d'onde (en nm)
n = np.array([1.77610, 1.76325, 1.74957, 1.73481, 1.72942, 1.72397]) # indices optiques
u_n = np.array([1.4, 1.4, 1.4, 1.3, 1.3, 1.3])*1e-4 # incertitudes-types sur les indices

# Ajustement affine selon une loi de Cauchy
p = np.polyfit(1/lamb**2, n, 1)
# p est un tableau contenant les coefficients de l'ajustement : p[0] pente ; p[1] ordonnée à l'origine

# Calcul des résidus et des écarts normalisés
n_mod = p[0]*1/lamb**2 + p[1] # calcul des ordonnées des points "modèle"
residus = n - n_mod # calcul des résidus
En = residus/u_n # calcul des écarts normalisés

# Vérification graphique
plt.figure(figsize = (19,5)) # création d'une fenêtre graphique avec trois graphes

plt.subplot(1, 3, 1) # 1er graphe
# points expérimentaux avec barres d'erreur pour n
plt.errorbar(1/lamb**2, n, yerr = u_n, fmt = 'bo', label = "Points expérimentaux")
# points issus de la régression linéaire
plt.plot(1/lamb**2, n_mod, 'c--', label = "Modèle affine")
plt.xlabel(r"$1/\lambda^2$ (en nm$^{-2}$)")
plt.ylabel(r"$n$")
plt.grid(), plt.legend(loc = 'best')

plt.subplot(1, 3, 2) # 2ème graphe
# Affichage des résidus avec barres d'incertitude-type :
plt.errorbar(lamb, residus, yerr = u_n, fmt = 'bo')
# Pour mieux visualiser la droite correspondant à un résidu nul :
plt.plot([350, 650], [0, 0], 'c--')
plt.xlabel(r"$\lambda$ (en nm)"), plt.xlim(350, 650)
plt.ylabel(r"$résidus$"), plt.ylim(-3e-4, 3e-4),
plt.ticklabel_format(axis = 'y', style = 'sci', scilimits = (0,0))
plt.grid()

plt.subplot(1, 3, 3) # 3ème graphe
# Affichage des écarts normalisés :
plt.plot(lamb, En, 'bo')
# Pour mieux visualiser le domaine des En acceptables :
plt.fill_between([350, 650], y1 = -2, y2 = 2, color = 'cyan', alpha = .1)
plt.xlabel(r"$\lambda$ (en nm)"), plt.xlim(350, 650)
plt.ylabel(r"$écarts normalisés$"), plt.ylim(-3,3)
plt.grid()
plt.show()

```

c) Moyenne et Incertitude-type

Pour déterminer la moyenne et l'incertitude-type sur les coefficients de Cauchy on va effectuer une simulation de Monte-Carlo.

```

# Simulation Monte-Carlo

## Nombre d'expériences simulées
N = 50000

## Initialisation des listes dans lesquelles on va stocker les valeurs de A et B correspondant à chaque expérience simulée
Amc = []
Bmc = []

## Pour chaque expérience, on simule la mesure des n pour les 6 longueurs d'onde
for i in range(N):
    # on tire aléatoirement (distribution normale ou uniforme) une série de valeurs de n
    nmc = n + rd.normal(0, u_n, size = 6)
    # on reprend l'ajustement affine avec cette série de valeurs de n
    p = np.polyfit(1/lamb**2, nmc, 1)
    # on stocke les valeurs des paramètres d'ajustement dans les listes AMC et BMC
    Amc.append(p[1])
    Bmc.append(p[0])

## Calcul de la valeur moyenne et de l'écart-type de A
Amoy = np.mean(Amc)
u_A = np.std(Amc, ddof = 1)
print("Estimation du paramètre A de la loi de Cauchy :")
print("* Valeur mesurée : A = {}".format(Amoy))
print("* Incertitude-type : u(A) = {}".format(u_A))
## Calcul de la valeur moyenne et de l'écart-type de B
Bmoy = np.mean(Bmc)
u_B = np.std(Bmc, ddof = 1)
print("Estimation du paramètre B de la loi de Cauchy :")
print("* Valeur mesurée : B = {}".format(Bmoy))
print("* Incertitude-type : u(B) = {}".format(u_B))
print("Attention à ne pas oublier les unités s'il y en a!")

Estimation du paramètre A de la loi de Cauchy :
* Valeur mesurée : A = 1.6844416364952528
* Incertitude-type : u(A) = 0.00018870742224982366

Estimation du paramètre B de la loi de Cauchy :
* Valeur mesurée : B = 14998.484514847492
* Incertitude-type : u(B) = 44.6579229323835
Attention à ne pas oublier les unités s'il y en a!

```

Ainsi on peut conclure que le modèle $n = A + \frac{B}{\lambda^2}$ est validé et on obtient :

$$\begin{cases} A = 1,68444 \pm 0,00019 \\ B = 14998 \pm 45 \text{ nm}^{-2} \end{cases}$$