

Mesures et incertitudes en sciences physiques

PLAN DU CHAPITRE

I	Mesure et variabilité de la mesure d'une grandeur physique	3
I.1	Position du problème	3
I.2	Fonction de densité de probabilité d'une variable aléatoire	4
I.3	Moyenne et incertitude-type	5
II	Grandeurs physiques mesurées directement	6
II.1	Bon Estimateur de la valeur du mesurande : la moyenne empirique	6
II.2	Bons estimateurs de l'incertitude-type	6
	a - Incertitude par évaluation statistique de type "A" ou "de répétabilité"	7
	b - Incertitude-type par évaluation probabiliste, dite "de type B"	8
III	Ecriture des résultats expérimentaux	9
III.1	Principe	9
III.2	Exemples	9
IV	Cas des grandeurs physiques calculées : propagation des incertitudes	10
IV.1	Estimateurs des valeurs de Y et $u(Y)$ par méthode analytique LPU : approximation linéaire	10
	a - Principe de calcul	10
	b - Expressions de $u(Y)$ dans quelques cas classiques (à connaître)	11
IV.2	Estimateurs des valeurs de Y et $u(Y)$ par simulation de type Monte-Carlo	12
	a - Principe d'évaluation	12
	b - Exemples	12
V	Comparaison d'une mesure à une valeur de référence : écart normalisé et z-score	14
V.1	Définition - critère de validité	14

V.2	Exemple	15
VI	Comparaison théorie-expérience - régression linéaire	17
VI.1	Régression affine - méthode des moindres carrés	17
	a - Paramètres idéaux (a,b) de la modélisation affine : ajustement au moindres carrés - calcul par Python	17
	b - Exemple complet de mise en oeuvre	18

I Mesure et variabilité de la mesure d'une grandeur physique

I.1 Position du problème

On peut définir l'action de mesurer une grandeur physique de la manière suivante :

Définition I-1: MESURE

Mesurer une grandeur physique X , nommée **mesurande**, vise à caractériser celle-ci quantitativement en lui attribuant une **valeur numérique x** , appelée **mesure** ou **estimation** ; celle-ci est toujours obtenue par comparaison avec une unité de valeur de même nature (longueur, temps, masse...), définie à partir d'un objet, d'une procédure, d'un dispositif, d'un matériau, dit **étalon** et reconnu officiellement.

Exemples :

- ▶ le **mètre étalon** du système international est défini depuis 1983 par le Bureau International des Poids et Mesures (BIPM) comme la distance parcourue par la lumière pendant une durée de 299 792 458^{ième} de seconde. Un dispositif gradué certifié, c'est à dire étalonné par un laboratoire spécialisé, par exemple un réglet ou un vernier, permet ensuite de réaliser des mesures de longueurs fiables.
- ▶ la **seconde étalon** du système international est définie comme la durée de 9 192 631 770 périodes de la radiation correspondant à la transition entre les deux niveaux hyperfins de l'état fondamental de l'atome de césium 133 non perturbé. Un chronomètre de qualité est donc également étalonné par un laboratoire spécialisé à partir d'une valeur de référence officielle (elle-même construite à partir de l'étalon de la seconde).

L'évaluation expérimentale du mesurande X présente toujours des "imperfections" qui interdisent l'accès à la valeur exacte de celui-ci ; on parle alors de **variabilité de la mesure** de cette grandeur. Les causes possibles de cette variabilité sont très nombreuses :

- soit **systématiques**, c'est à dire apparaissant à chaque mesure, et provenant généralement d'une dérive de l'instrument de mesure ou une imperfection de protocole, par exemple :
 - ◊ mauvaise position de l'origine de graduation angulaire de la lunette de visée d'un goniomètre (on procèdera alors à une double mesure pour juguler cette imperfection !).
 - ◊ instrument électronique "âgé" entraînant une dérive systématique de la mesure ou bien une puce employée dans celui-ci et présentant une anomalie ; par exemple, une mesure de tension avec deux voltmètres peut donner des résultats légèrement différents.
 - ◊ ...
- soit **aléatoires** par exemple :
 - ◊ liées à l'opérateur qui par exemple commet des erreurs : par exemple une erreur de parallaxe lors de l'emploi d'une lunette de visée.
 - ◊ liées à la précision intrinsèque de l'appareil exploité : mesure d'un angle sur la graduation d'un goniomètre précis au mieux à 1' d'angle
 - ◊ liées aux conditions de mesures ; par exemple, la vitesse de son dépendant de la température, des mesures à différents moments de la journée donneront des résultats légèrement différents...

Les causes de la composante systématique peuvent parfois être éliminées, par exemple en changeant d'appareil ou de protocole ; en revanche, les **composantes aléatoires** seront toujours présentes lors d'un

mesurage et entraîne une dispersion des valeurs mesurées autour d'une valeur exacte inaccessible.

Définition I-2: INCERTITUDE DE MESURE

Le BIPM définit depuis 1993 l'incertitude de mesure comme : «un paramètre, associé au résultat d'un mesurage, qui caractérise la dispersion des valeurs "qui pourraient raisonnablement" être attribuées au mesurande».

L'incertitude $u(X)$ sur le mesurande X , dite **incertitude-type** ("u" pour *uncertainty*¹) permet ainsi d'informer sur le **niveau de fiabilité de la valeur estimée de X** ,

IMPORTANT : le calcul des incertitudes sur toutes les grandeurs expérimentales est évidemment attendu dans la rédaction de vos comptes-rendus de TP.

I.2 Fonction de densité de probabilité d'une variable aléatoire

Une grandeur physique X dont la valeur évolue de manière continue est caractérisée par une fonction de densité de probabilité $\varphi(x)$:

Définition I-3: FONCTION DE DENSITÉ DE PROBABILITÉ

On appelle fonction de densité de probabilité de la grandeur X , la fonction $\varphi(x)$ telle que la probabilité que la valeur x de X se trouve dans l'intervalle $[a,b]$ s'écrit :

$$P(a \leq x \leq b) = \int_a^b \varphi(x) \cdot dx$$

Ainsi, la probabilité élémentaire que la mesure de X se trouve entre x et $x + dx$ s'écrit :

$$dP(x) = \varphi(x) \cdot dx$$

La fonction de densité de probabilité est évidemment normalisée : $\int_{\mathbb{R}} \varphi(x) \cdot dx = 1$

Deux fonctions de distribution de probabilité jouent un rôle important dans la mesure physique :

- la fonction de distribution gaussienne ou loi normale :

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Cette distribution dépend de deux paramètres : $\mu \in \mathbb{R}$, et $\sigma \in \mathbb{R}^+$.

1. *incertitude en anglais*

- la fonction de distribution uniforme (ou "loi du pire") :

$$\begin{cases} \varphi(x) = \frac{1}{2\Delta} \text{ si } \mu - \Delta \leq x \leq \mu + \Delta \\ \varphi(x) = 0 \text{ sinon} \end{cases}$$

Cette distribution dépend de deux paramètres : $\mu \in \mathbb{R}$, et $\Delta \in \mathbb{R}^+$.

I.3 Moyenne et incertitude-type

A partir de la loi de probabilité $\varphi(x)$ d'une grandeur physique X , on définit sa moyenne et son incertitude type :

Définition I-4: ESPÉRANCE (MOYENNE "VRAIE") ET INCERTITUDE-TYPE

On définit pour une grandeur X de fonction de densité de probabilité $\varphi(x)$:

► la moyenne "vraie" comme l'espérance mathématique de la grandeur X : $\bar{x} = E(X) = \int_{\mathbb{R}} x \cdot \varphi(x) \cdot dx$

► l'incertitude type ou écart-type des valeurs comme la racine carrée de la variance : $u(X) = \sigma_x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{\overline{(x - \bar{x})^2}} = \sqrt{\int_{\mathbb{R}} (x - \bar{x})^2 \cdot \varphi(x) \cdot dx}$

EXPRESSION POUR LES LOIS "CLASSIQUES" :

- Distribution uniforme :

$$\bar{x} = \int_{-\infty}^{+\infty} x \cdot \varphi(x) \cdot dx = \frac{1}{2\Delta} \int_{\mu-\Delta}^{\mu+\Delta} x \cdot dx = \frac{1}{4\Delta} ((\mu + \Delta)^2 - (\mu - \Delta)^2) \implies \boxed{\bar{x} = \mu}$$

$$\begin{aligned} u(X)^2 = \sigma^2 &= \int_{-\infty}^{+\infty} (x - \bar{x})^2 \cdot \varphi(x) \cdot dx = \frac{1}{2\Delta} \int_{\mu-\Delta}^{\mu+\Delta} (x - \bar{x})^2 \cdot dx \stackrel{u=x-\bar{x}}{=} \frac{1}{2\Delta} \int_{\mu-\Delta}^{\mu-\Delta} u^2 \cdot du = \frac{1}{3} \frac{1}{\Delta} \Delta^3 \\ &= \frac{\Delta^2}{3} \quad (\text{I.1}) \end{aligned}$$

soit : $\boxed{u(X) = \sigma_x = \frac{\Delta}{\sqrt{3}}}$

- Distribution normale :

$\boxed{\bar{x} = \mu}$ de manière évidente, et on admettra que $\boxed{u(X) = \sigma_x = \sigma}$

II Grandeurs physiques mesurées directement

II.1 Bon Estimateur de la valeur du mesurande : la moyenne empirique

Si l'on procède à N mesures successives x_1, x_2, \dots, x_N d'une grandeur X , il est communément admis de retenir la **moyenne "empirique"** de cet échantillon comme bon estimateur de la valeur de X , c'est à dire la moyenne arithmétique sur l'échantillon. On peut en effet dégager la pertinence de ce choix par la méthode des moindres carrés :

On recherche la valeur de x qui minimise la somme des carrés des écarts entre chaque valeur mesurée x_k et la valeur recherchée x :

$$\frac{\partial}{\partial x} \sum_{k=1}^N (x_k - x)^2 = -2 \sum_{k=1}^N (x_k - x_{est}) = -2 \sum_{k=1}^N x_k + 2Nx_{est} = 0$$

soit :

$$x_{est} = x_{emp} = \frac{1}{N} \sum_{k=1}^N x_k$$

(NB : un calcul immédiat de la dérivée seconde montre que celle-ci est positive ce qui confirme bien que le résultat recherché correspond à un minimum de la somme des carrés.)

remarque 1 : la moyenne empirique x_{emp} est une variable aléatoire répartie selon une loi normale centrée sur l'espérance $E(X) = \bar{x}$, **valeur inaccessible**.

remarque 2 : Pour améliorer l'estimation, il est communément admis qu'il faut augmenter le nombre N de mesures ; c'est un résultat statistique que l'on démontre assez facilement (cf document Incertitudes disponible sur le site).

A RETENIR :

Propriété II-1: ESTIMATEUR DU MESURANDE

Le meilleur estimateur d'une grandeur X mesurée à partir d'un N-échantillon de mesures $\{x_1, x_2, \dots, x_N\}$ est obtenu en faisant la moyenne empirique de celui-ci :

$$x_{emp} = \frac{1}{N} \sum_{k=1}^N x_k$$

II.2 Bons estimateurs de l'incertitude-type

a - Incertitude par évaluation statistique de type "A" ou "de répétabilité"

Si l'on procède à N mesures de la grandeur X x_1, x_2, \dots, x_N alors on montre que² le meilleur estimateur de l'incertitude-type sur une mesure de la grandeur X est :

$$u(X) \simeq \sigma(X)_{N-1} = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (x_k - x_{emp})^2} = \sqrt{\frac{1}{N-1} \left[\sum_{k=1}^N x_k^2 - \frac{1}{N} \left(\sum_{k=1}^N x_k \right)^2 \right]}$$

Par ailleurs, on montre immédiatement (hypothèse d'indépendance des N mesures) que : $u(x_{emp}) = \frac{1}{\sqrt{N}} u(X) \simeq$ donc :

$$u(x_{emp}) = \sigma_{x_{emp}} \simeq \frac{\sigma(X)_{N-1}}{\sqrt{N}}$$

On constate ici un résultat assez intuitif : augmenter le nombre de mesures améliore sensiblement la valeur de l'estimateur $\sigma_{x_{emp}}$.

EXEMPLE D'EXPLOITATION :

Supposons que l'on dispose d'une série de 10 évaluations de la distance focale f' d'une lentille par autocollimation :

Groupe	1	2	3	4	5	6	7	8	9	10
$f' \text{ (cm)}$	20,1	19,1	19,9	20,4	20,2	19,7	19,6	18,2	19,4	20,3

Le script python suivant évalue les estimateurs f_{emp} de la moyenne, de l'incertitude-type sur la valeur de f , et enfin l'incertitude-type sur f_{emp} :

Listing I.1 –

```
1 import numpy as np
2 f=np.array([20.1, 19.1, 19.9, 20.4, 20.2, 19.7, 19.6, 18.2, 19.4, 20.3], dtype=
    float)
3 N=len(f)
4 femp=np.mean(f)
5 uf=np.std(f, ddof=1) # le paramètre ddof=1 permet le calcul de l'écart type avec N-1
    au dénominateur
6 print("Moyenne empirique des valeurs : femp=", femp)
7 print(r"Incertitude-type associée à f : u(f)={0}\{1}".format(uf, "cm"))
8 print(r"Incertitude-type associée à femp : u(femp)={0}\{1}".format(uf/np.sqrt(N), "cm"
    ))
```

Qui donne la sortie suivante :

In [1] :

Moyenne empirique des valeurs : femp= 19.69 cm

Incertitude-type associée à f : u(f)=0.6674162452656096 cm

Incertitude-type associée à femp : u(femp)=0.211055482436897 cm

2. cf doc. "Incertitudes" sur le site

b - Incertitude-type par évaluation probabiliste, dite "de type B"

La méthode d'évaluation statistique des incertitudes (type A) nécessite un nombre suffisant de mesures de la grandeur étudiée. Cette approche n'est parfois pas envisageable pratiquement, soit parce qu'on ne dispose pas du temps nécessaire pour le faire, ou bien parce que la collecte du N-échantillon de mesures (avec par exemple $N \geq 10$) se fait par un processus trop coûteux (par exemple un essai à rupture d'une pièce).

Dans ces conditions, on ne réalise qu'une seule mesure de la grandeur X , et c'est la connaissance d'informations diverses sur les causes de variabilité de la mesure de X qui permet alors d'évaluer l'incertitude :

- **Résultats antérieurs de mesures** dans les mêmes conditions ; par exemple un autre groupe de TP qui aurait réalisé la manipulation exactement avec le même matériel et procédé à une évaluation de type A pour $u(X)$! Rare, mais pourquoi pas !)
- **Caractéristiques de dispersion des appareils de mesures**, attestées par le constructeur (en conformité avec la norme ISO 10012 :2003 si les choses sont bien faites).

EXEMPLE : la mesure de la valeur efficace d'un signal sinusoïdal pur de fréquence 14,8 kHz donne $U_{eff} = 4,785 \text{ V}$

Les caractéristiques du voltmètre employé (METRIX MX5060) en mode AC sont les suivantes :



Gamme	Domaine de fonctionnement	Domaine de mesure spécifié ³⁾	Résolution	Incertitude (\pm)
600 mV	0 à 600,0 mV	60,0 à 600,0 mV	0,1 mV	1 % L + 0,25 % x [F(kHz)-1]L ± 5 D
6 V	0 à 6,000 V	0,600 à 6,000 V	0,001 V	1 % L + 0,18 % x [F(kHz)-1]L ± 3 D
60 V	0 à 60,00 V	6,00 à 60,00 V	0,01 V	
600 V	0 à 600,0 V	60,0 à 600,0 V	0,1 V	
1000 V ²⁾	0 à 1000 V	60 à 1000 V	1 V	

On constate qu'à partir de la valeur du calibre (gamme) employé, la mesure effectuée L , la fréquence du signal F , et enfin la plus petite valeur affichée sur le calibre choisi, on peut calculer la demi-largeur Δ de l'intervalle de mesure de l'appareil. Dans le cas présent cela donne (ligne "gamme 6V" du tableau) :

$$\Delta = \frac{1}{100} \times U_{eff} + \frac{0,18}{100} \times (14,8 - 1) \times U_{eff} + 3 \times 1.10^{-3} = 0,0086874 \text{ V}$$

On supposera ici que toutes les mesures dans l'intervalle $[U_{eff} - \Delta, U_{eff} + \Delta]$ sont équiprobables, ce qui est ici l'hypothèse la plus raisonnable compte tenu des informations dont on dispose ; la distribution de probabilité des valeurs est donc **uniforme** sur cet intervalle, et nous avons donc pour l'incertitude :

$$u(U_{eff}) = \frac{2\Delta}{\sqrt{12}} = \frac{\Delta}{\sqrt{3}} = \frac{0,0086874}{\sqrt{3}} = 0,0050156 \simeq 5,0 \text{ mV}$$

- **Précision d'une lecture réalisée par l'opérateur**, par exemple sur un système gradué pour lequel on est certain que la valeur se trouve entre deux "traits" de graduation, ou bien deux positions connues.

EXEMPLE :

On mesure une position angulaire θ sur la graduation d'un goniomètre précis à 1' d'angle ; on estime dans ce cas que la lecture de la valeur peut se faire **au mieux entre deux graduations** θ_{inf} et θ_{sup} séparées d'1' d'angle (2Δ), et que toutes les valeurs réparties entre ces deux graduations sont équiprobables ; cela revient là-encore à choisir une **distribution de probabilité uniforme** entre θ_{inf} et θ_{sup} , et conduit à une incertitude sur θ :



$$u(\theta) = \frac{\theta_{sup} - \theta_{inf}}{\sqrt{12}} = \frac{1}{60} \cdot \frac{2\Delta}{2\sqrt{3}} = 4,811 \cdot 10^{-3} \text{ }^\circ \simeq 5 \cdot 10^{-3} \text{ }^\circ$$

III Ecriture des résultats expérimentaux

III.1 Principe

On retiendra les règles simples suivantes pour écrire le résultat d'une mesure x de la grandeur X :

- on limitera l'écriture de l'incertitude-type $u(X)$ avec **deux** chiffres significatifs.
- on écrira l'incertitude-type $u(X)$ dans la même unité et avec la même puissance de 10 que la valeur x
- on écrira enfin la valeur de x avec un nombre de chiffres après la virgule **identique** à celui de $u(X)$

III.2 Exemples

On reprend les exemples vus plus hauts :

- MESURE D'UNE DISTANCE FOCALE PAR AUTOCOLLIMATION (ÉVALUATION $u(X)$ TYPE A) : on avait $x = 19,69 \text{ cm}$ et $u(X) = 0,66741624 \text{ cm}$, donc on écrira :

$$f' = (19,69 \pm 0.67) \text{ cm}$$

- MESURE D'UNE TENSION (ÉVALUATION $u(X)$ TYPE B) :
on avait $U_{eff} = 4,785 \text{ V}$ et $u(X) = 5,0 \text{ mV}$ =, donc on écrira :

$$U_{eff} = (4,7850 \pm 0,0050) \text{ V}$$

IV Cas des grandeurs physiques calculées : propagation des incertitudes

Lorsque la grandeur à évaluer Y n'est pas directement accessible par la mesure mais liée aux mesurandes (X_1, X_2, \dots, X_n) par une relation :

$$Y = F(X_1, X_2, \dots, X_n)$$

l'estimation de Y et celle de son incertitude-type $u(Y)$ peuvent se faire par deux méthodes distinctes :

- **méthode analytique** dite LPU pour Law of propagation of Uncertainty³ ou **méthode de propagation des incertitudes (ou des variances)**, dont l'usage doit être limité au cas de fonctions F très simples.
- **méthode par simulation type «Monte-Carlo»** qui est désormais la recommandation du BIPM ; c'est la méthode à privilégier lorsque la fonction F est compliquée, au point de rendre délicate (voire fausse) si fortes non linéarités et les $u(x_i)$ élevés) l'évaluation de $u(Y)$.

IV.1 Estimateurs des valeurs de Y et $u(Y)$ par méthode analytique LPU : approximation linéaire

a - Principe de calcul

Si les incertitudes sur les variables x_1, x_2, \dots sont faibles, on peut raisonnablement poser³ que l'estimateur de Y s'obtient par une approximation linéaire avec :

$$y_{emp} \simeq F(x_{1emp}, x_{2emp}, \dots, x_{nemp})$$

A RETENIR :

Propriété IV-1: INCERTITUDE PROPAGÉE PAR MÉTHODE LPU

Un estimateur y_{emp} de la grandeur Y reliée aux mesurandes X_i par la relation $Y = F(X_1, X_2, \dots, X_n)$ est donnée par :

$$y_{emp} = F(x_{1emp}, x_{2emp}, \dots, x_{nemp})$$

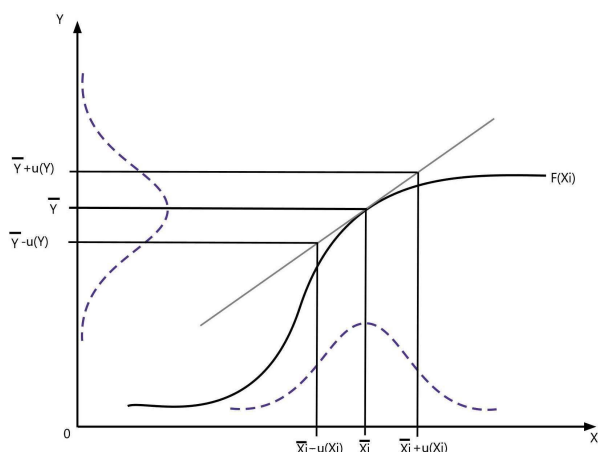
L'incertitude sur Y est donnée par :

$$u(Y) = \sqrt{\sum_{i=1}^n \left(\frac{\partial F}{\partial x_i} \right)^2 u^2(X_i)}$$

les $u(X_i)$ étant les incertitudes-type connues sur les grandeurs X_i .

Le graphe ci-dessous résume assez simplement le principe de propagation des incertitudes $u(X_i)$ sur la grandeur Y dans l'approximation linéaire :

3. cf démonstration dans doc. "Incertitudes" sur le site

FIGURE I.1 – Illustration graphique de la propagation des incertitudes $U(X_i)$ par approximation linéaire

NB : dans le cas d'une relation F linéaire par rapport à chacune de ses variables, la relation de propagation des incertitudes est rigoureusement valable et simple à exploiter.

b - Expressions de $u(Y)$ dans quelques cas classiques (à connaître)

Le tableau ci-dessous rassemble quelques exemples classiques de propagation d'incertitudes :

$Y = F(X_i)$	$u(Y)$	Exemples
$y = a \pm b$	$u(Y) = \sqrt{u^2(A) + u^2(B)}$	Déviation angulaire $D = \theta_2 - \theta_1 \Rightarrow u(D) = \sqrt{u^2(\theta_1) + u^2(\theta_2)}$
$y = a \times b$	$u(Y) = y \times \sqrt{\frac{u^2(A)}{a^2} + \frac{u^2(B)}{b^2}}$	Puissance électrique $P = U \times I \Rightarrow u(P) = P \times \sqrt{\frac{u^2(U)}{U^2} + \frac{u^2(I)}{I^2}}$
$y = \frac{a}{b}$	$u(Y) = y \times \sqrt{\frac{u^2(A)}{a^2} + \frac{u^2(B)}{b^2}}$	Résistance électrique $R = \frac{U}{I} \Rightarrow u(R) = R \times \sqrt{\frac{u^2(U)}{U^2} + \frac{u^2(I)}{I^2}}$
$y = a^\alpha \times b^\beta$	$u(Y) = y \times \sqrt{\alpha^2 \frac{u^2(A)}{a^2} + \beta^2 \frac{u^2(B)}{b^2}}$	Puissance (dans un résistor) $P = \frac{U^2}{R} \Rightarrow u(P) = P \times \sqrt{4 \frac{u^2(U)}{U^2} + \frac{u^2(R)}{R^2}}$

EXEMPLE :

Pour mesurer la distance algébrique lentille-objet \overline{OA} sur un banc d'optique précis au millimètre, on effectue **deux mesures de positions** (celle du point A et celle du point O) :

$$\overline{OA} = x_A - x_O \Rightarrow u(\overline{OA}) = \sqrt{u^2(x_A) + u^2(x_O)}$$

$$\text{avec : } u(x_A) = u(x_O) = \frac{1.10^{-3}}{\sqrt{12}} = 0,00029 \text{ m} = 0,029 \text{ cm}$$



On mesure par exemple $\overline{OA}_{est} = -41,3 \text{ cm}$

$$\text{A.N. : } u(\overline{OA}) = \sqrt{\left(\frac{10^{-3}}{\sqrt{12}}\right)^2 + \left(\frac{10^{-3}}{\sqrt{12}}\right)^2} = 0,00041 \text{ m} \simeq 0,041 \text{ cm} \text{ soit : } \frac{u(\overline{OA})}{\overline{OA}} \simeq 0,1\%$$

Le résultat final s'écrira donc : $\boxed{\overline{OA} = (-41,300 \pm 0,041) \text{ cm}}$

IV.2 Estimateurs des valeurs de Y et $u(Y)$ par simulation de type Monte-Carlo

a - Principe d'évaluation

Dans le cas d'une fonction F engendrant des calculs d'incertitudes trop complexes, ou bien si celle-ci présente une non linéarité vis à vis de variables possédant une forte incertitude-type, on a recours à la méthode de simulation de Monte-Carlo, très simple à mettre en oeuvre pour évaluer la propagation des incertitudes.

PRINCIPE :

On rappelle que l'on cherche une estimation de la grandeur calculée $Y = f(X_1, X_2, \dots, X_n)$ et son incertitude-type $u(Y)$.

- On procède à un nombre important de tirages aléatoires, par exemple $N = 10^5$, de chaque grandeur X_i $i \in [1, n]$ selon une loi de densité de probabilité choisie et supposée valable pour cette grandeur (concrètement, on se limite aux distributions **uniforme** et **normale**). On dispose alors de N jeux de valeurs (x_1, x_2, \dots, x_n) .
- On calcule pour chaque jeu $(x_1, x_2, \dots, x_n)_k$ avec $k \in [0, N-1]$ la valeur de $y_k = Y((x_1, x_2, \dots, x_n)_k)$. On dispose donc d'un jeu de N valeurs de la grandeur Y .
- On en déduit les estimations des valeurs de Y et $u(Y)$ en utilisant les "bons" estimateurs vus plus haut, c'est à dire la moyenne empirique de Y sur les N valeurs, et l'écart type débiaisé à savoir :

$$\boxed{\begin{aligned} \bar{y} &\simeq y_{emp} = \frac{1}{N} \sum_{k=0}^{N-1} y_k \\ u(Y) &\simeq \sigma(Y)_{N-1} = \sqrt{\frac{1}{N-1} \sum_{k=0}^{N-1} (y_k - y_{emp})^2} \end{aligned}}$$

b - Exemples

● CALCUL D'UNE DISTANCE ALGÈBRE SUR UN BANC D'OPTIQUE :

on reprend ici la mesure de la distance \overline{OA} sur le banc d'optique vu plus haut :

$$\overline{OA} = x_A - x_O$$

On supposera : $x_A = -58,42 \text{ cm}$, $x_O = -17,12 \text{ cm}$, et avec x_A et x_O distribués selon une loi uniforme avec une demi-largeur $\Delta = 0,5 \text{ mm}$.

Pour le script python, on peut s'appuyer sur les possibilités des tableaux du module `numpy` qui permettent de faire des calculs terme à terme sur les éléments des tableaux lorsque ceux-ci sont de même dimension :

Listing I.2 – Méthode Monte-Carlo (exemple 1)

```

1 import numpy as np
2 import numpy.random as rd
3
4 ##### Données numériques #####
5 N=1e5
6 xA=-58.42
7 xO=-17.12
8 deltaxA=0.5e-1
9 deltaxO=0.5e-1
10
11 ##### Construction des tableaux de tirages #####
12 tabxA=rd.uniform(xA-deltaxA,xA+deltaxA,N)
13 tabxO=rd.uniform(xO-deltaxO,xO+deltaxO,N)
14
15 ##### Construction du tableau de OA #####
16 OA=tabxA-tabxO
17
18 ##### Estimateurs de OA et de l'incertitude #####
19 OAemp=np.mean(OA)
20 uOA=np.std(OA, ddof=1)
21 print(r"La moyenne empirique est :  $\overline{OA}_{\text{emp}} = \{0\}_{1}$ ".format(OAemp, "cm"
22 ))
21 print(r"L'incertitude-type est :  $u(\overline{OA}) = \{0\}_{1}$ ".format(uOA, "cm"))

```

qui donne :

In [1]: La moyenne empirique est -41.30014459275822 cm

L'incertitude-type est 0.04086800079834664 cm

résultat conforme à ce que nous avait renvoyé le calcul par méthode LPU.

● CALCUL DE L'INDICE n DU MATÉRIAU D'UN PRISME.

Considérons un prisme d'angle au sommet A éclairé par un faisceau fin de lumière en situation de minimum

de déviation D_m . On montre que l'on a dans ces conditions :

$$\left[\begin{array}{l} i_m = i'_m \\ r_m = r'_m = \frac{A}{2} \\ D_m = 2i_m - A \Rightarrow i_m = \frac{D_m + A}{2} \end{array} \right.$$

La loi de Descartes : $n \cdot \sin r_m = \sin i_m$

permet de dégager l'expression de l'indice du matériau du prisme :

$$n = \frac{\sin\left(\frac{D_m + A}{2}\right)}{\sin \frac{A}{2}}$$

L'estimation de n et de son incertitude est une très bonne illustration de la puissance de la méthode Monte-Carlo. Il n'y a ici que deux mesurandes, l'angle au sommet du prisme A et l'angle de déviation minimale D_m , mais le calcul de $u(n)$ par méthode LPU est lourd pour une simple évaluation d'incertitude ; on donne ici le résultat de ce calcul qui est assez parlant ! :

$$u(n) = n \sqrt{\frac{1}{4} \cotan^2 \left(\frac{D_m + A}{2} \right) \cdot u^2(D_m) + \frac{1}{4} \left[\cotan \left(\frac{D_m + A}{2} \right) - \cotan \frac{A}{2} \right]^2 \cdot u^2(A)}$$

On supposera les valeurs mesurées expérimentalement : $A = 60^\circ$ et $D_m = 37,2^\circ$, toutes les deux distribuées selon une loi uniforme avec la demi-largeur $\Delta = 0,5^\circ$

Le calcul théorique donne alors $n = 1,50022$ et $u(n) = 0,004629$.

On va maintenant déterminer n et son incertitude type $u(n)$ par méthode MC, ce que réalise le code python ci-dessous :

Listing I.3 –

```

1 import numpy as np
2 import numpy.random as rd
3
4 ##### Données numériques #####
5 pi=np.pi
6 A=60*pi/180
7 deltaA=0.5*pi/180 #demi-largeur en radian de la distribution uniforme pour A
8 Dm=37.2*pi/180
9 deltaDm=0.5*pi/180 #demi-largeur en radian de la distribution uniforme pour Dm
10 N=10000
11
12 ##### Construction des tableaux de tirages #####
13 tabA=rd.uniform(A-deltaA,A+deltaA,N)
14 tabDm=rd.uniform(Dm-deltaDm,Dm+deltaDm,N)
15
16 ##### Construction du tableau de valeurs de n #####
17 tabn=np.sin((tabDm+tabA)/2)/np.sin(tabA/2)
18
19 ##### Estimateurs de n et de l'incertitude u(n) #####
20 n_emp=np.mean(tabn)
21 u_n=np.std(tabn,ddof=1)
22 print(r"La moyenne empirique de l'indice est : n=",n_emp)
23 print(r"L'incertitude-type sur l'indice est u(n)=",u_n)

```

Là encore, les résultats sont conformes à ceux obtenus par méthode LPU :

In [7]: La moyenne empirique de l'indice est: n= 1.5002303091323923

L'incertitude-type sur l'indice est u(n)= 0.004599468877324797

V Comparaison d'une mesure à une valeur de référence : écart normalisé et z-score

V.1 Définition - critère de validité

Supposons que l'on réalise la mesure x d'une grandeur X dont on connaît une valeur de référence $x_{\text{réf}}$, donnée par exemple par un calcul théorique et supposée sans incertitude ; si cette mesure est de bonne qualité (bon

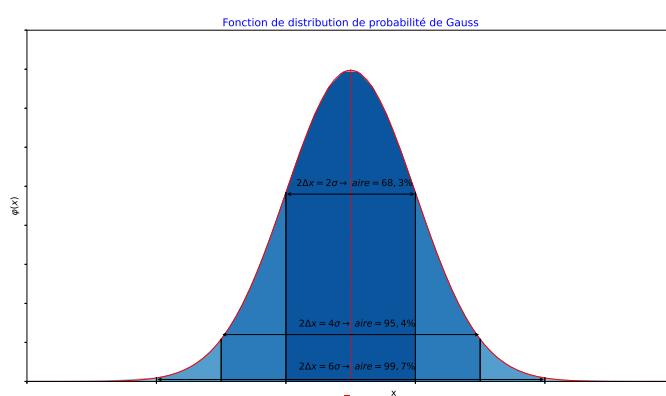
protocole employé, attention particulière de l'opérateur etc...), la valeur de x doit être compatible avec $x_{\text{réf}}$, en ce sens que leur écart doit être acceptable.

Définition V-1: Z-SCORE

Pour caractériser quantitativement l'écart entre x et $x_{\text{réf}}$, on introduit une nouvelle grandeur Z appelée écart normalisé ou z-score, définie par :

$$Z = \frac{X - x_{\text{réf}}}{u(X)}$$

Supposons par exemple que la grandeur x soit distribuée selon une loi normale ($u(x), x_{\text{réf}} = \bar{x}$) ; dans ce cas, $\simeq 95\%$ des mesures effectuées vont se trouver dans l'intervalle $[x_{\text{réf}} - 2u(X), x_{\text{réf}} + 2u(X)]$



Dans ces conditions, Z est une grandeur distribuée selon la loi normale telle que :

$$\begin{cases} \bar{Z} = \frac{\bar{x} - \bar{x}}{u(X)} = 0 \\ u(Z) = \frac{u(x - \bar{x})}{u(X)} = \frac{u(X)}{u(X)} = 1 \end{cases}$$

On retiendra la propriété suivante :

Propriété V-1: CRITÈRE DE COMPATIBILITÉ D'UNE GRANDEUR VS RÉFÉRENCE

Ainsi, pour que la mesure x soit qualifiée de "compatible" avec la valeur de référence choisie \bar{x} , c'est à dire avec 95% de chance d'être dans l'intervalle $[x_{\text{réf}} - 2u(X), x_{\text{réf}} + 2u(X)]$, il faut satisfaire le critère suivant :

$$z = |Z| = \frac{|x - x_{\text{réf}}|}{u(X)} \leq 2$$

V.2 Exemple

Reprenons ici le cas de la mesure d'une distance focale par autocollimation vue en II.2.a- ; on rappelle le tableau de valeurs :

Groupe	1	2	3	4	5	6	7	8	9	10
$f \text{ (cm)}$	20,1	19,1	19,9	20,4	20,2	19,7	19,6	18,2	19,4	20,3

On reprendra pour le calcul du z-score la valeur de l'incertitude-type estimée arrondie, soit : $u(f) = 0.2 \text{ cm}$ et l'on supposera que la valeur de référence est la moyenne empirique (un peu à tort car celle-ci est également une variable entachée d'une incertitude ! Mais il faut bien prendre une valeur en TP !).

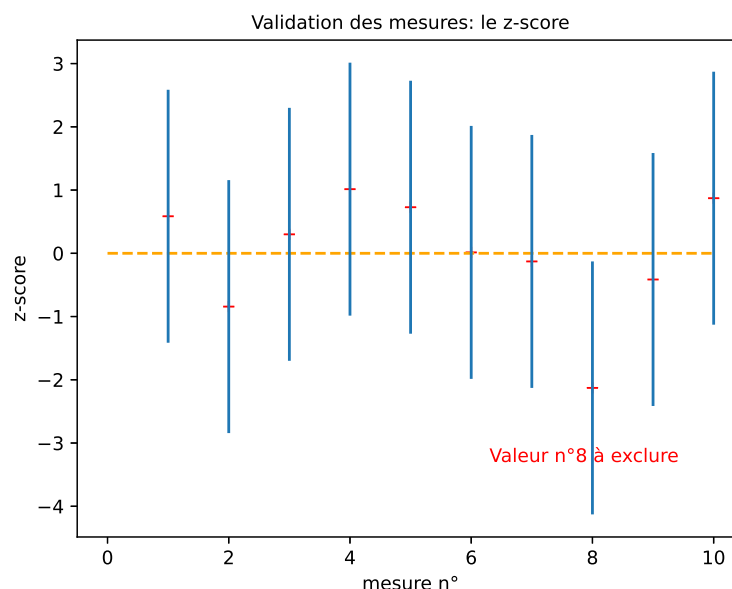
Listing I.4 – Tracé du z-score

```

1 import numpy as np
2 from matplotlib import pyplot as plt
3 n_mes=np.arange(1,11,1)
4 f=np.array([20.1, 19.1, 19.9, 20.4, 20.2, 19.7, 19.6, 18.2, 19.4, 20.3])
5 N=len(f)
6 fref=np.mean(f)
7 uf=round(np.std(f, ddof=1),1) # on arrondit uf à 1 décimale
8 Z=(f-fref)/uf
9
10 ##### Représentation graphique du z-score #####
11 plt.plot(n_mes,Z,c="red",marker='+',linestyle='_')
12 plt.plot([0, 10], [0, 0], c='orange', linestyle = '—')
13 plt.title("Validation des mesures: le z-score", fontsize=10)
14 plt.xlabel("mesure n°")
15 plt.ylabel("z-score")
16 plt.text(6.3, -3.3, "Valeur n°8 à exclure", color="red")
17 plt.errorbar(n_mes,Z,2, linestyle='_')
18 plt.show()

```

On obtient le tracé suivant sur lequel on constate que a mesure n°8 est à exclure :



NB : le tracé des barres d'erreurs de hauteur ± 2 sur chaque valeur expérimentale permet de rapidement vérifier sa validité.

IMPORTANT : si l'on réalise une seule mesure d'une grandeur, on la comparera à la valeur de référence **donnée par une théorie** ; il faudra alors s'assurer que le critère de compatibilité est bien vérifié. Par exemple, pour la mesure de la vitesse du son dont la valeur est connue pour T fixée :

$$c_{theo}(293\text{ K}) = (344,1 \pm 1) \text{ m.s}^{-1} \text{ et } c_{exp} = (345 \pm 4) \text{ m.s}^{-1}$$

$$\text{Le z-score vaut } z = \frac{c_{exp} - c_{theo}(293\text{ K})}{u(c)} = 0,225 \text{ ce qui valide la valeur } c_{exp}.$$

VI Comparaison théorie-expérience - régression linéaire

VI.1 Régression affine - méthode des moindres carrés

On dispose d'un ensemble de N valeurs expérimentales $(x, y) = (x_1, x_2, \dots, x_N ; y_1, y_2, \dots, y_N)$ des grandeurs physiques X et Y ; nous supposons un possible lien affine entre les deux, c'est à dire l'existence d'une relation de type $y = a \cdot x + b$.

La **régression affine** est une opération qui consiste à choisir au mieux les paramètres a et b pour que les valeurs expérimentales soient les plus "compatibles" possibles avec cette relation affine.

a - Paramètres idéaux (a,b) de la modélisation affine : ajustement au moindres carrés - calcul par Python

On se placera dans les hypothèses simples suivantes :

- l'incertitude-type sur X est négligeable
- les incertitudes sur chaque valeur de Y sont connues $u(Y_i)$

On peut montrer que le critère pour choisir la **meilleure droite affine**, et donc déterminer les meilleurs paramètres a et b , est de minimiser la quantité :

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

par rapport aux paramètres a et b , soit d'avoir :

$$\begin{cases} \frac{\partial D}{\partial a} = 0 \\ \frac{\partial D}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^N (y_i - ax_i - b)x_i = 0 \\ \sum_{i=1}^N (y_i - ax_i - b) = 0 \end{cases} \Rightarrow \begin{cases} a \sum_{i=1}^N x_i^2 + b \sum_{i=1}^N x_i - \sum_{i=1}^N (x_i y_i) = 0 \\ a \sum_{i=1}^N x_i + Nb - \sum_{i=1}^N y_i = 0 \end{cases}$$

qui conduit sans peine en divisant par N et en utilisant les moyennes "empiriques" :

$$\begin{cases} a(x^2)_{emp} + bx_{emp} - (xy)_{emp} = 0 \\ ax_{emp} + b - y_{emp} = 0 \end{cases}$$

On en tire sans peine les expressions de a et b : $a = \frac{(xy)_{emp} - x_{emp}y_{emp}}{(x^2)_{emp} - x_{emp}^2}$ et $b = y_{emp} - ax_{emp}$

La fonction python `polyfit` du module `numpy` permet de réaliser un ajustement au sens des moindres carrés d'une série de valeurs avec un polynôme, qui sera ici de degré 1. Elle renvoie sous forme d'un tableau les valeurs de a et b . La syntaxe de base (i.e. sans paramètres optionnels) est la suivante :

`numpy.polyfit(X, Y, 1)`

b - Exemple complet de mise en oeuvre

On suppose que lors d'une séance de travaux pratiques, un groupe d'élèves a réalisé à l'aide d'un logiciel la détermination expérimentale du temps caractéristique de charge d'un condensateur d'un circuit RC alimentée par un générateur de tension continue. L'incertitude sur τ donnée par le logiciel est consignée dans le tableau de valeurs.

$R \text{ (}\Omega\text{)}$	100	200	300	400	500	600	700	800	900	1000
$\tau \text{ (en ms)}$	151	240	356	458	547	651	741	843	949	1041
$u(\tau) \text{ (en ms)}$	6,04	9,6	14,24	18,32	21,88	26,04	29,64	33,73	37,96	41,64

On supposera par ailleurs négligeable l'incertitude sur les valeurs de résistances.

OBJECTIF : déterminer la capacité C du condensateur.

Le modèle théorique donne $\tau = RC$ donc du type $y = ax + b$ avec $b = 0$. Cette spécificité sera intéressante !

Le code python suivant permet le tracé expérimental $\tau = fct(R)$, celui de la droite de régression linéaire $y = aR + b$, des barres d'erreurs, les résidus normalisés définis par $\frac{\tau_i - (a \times R_i + b)}{u(\tau_i)}$ (écarts normalisés (type z-score) entre valeur expérimentale et théorique), et enfin l'affichage des paramètres a et b :

Listing I.5 – Régression linéaire

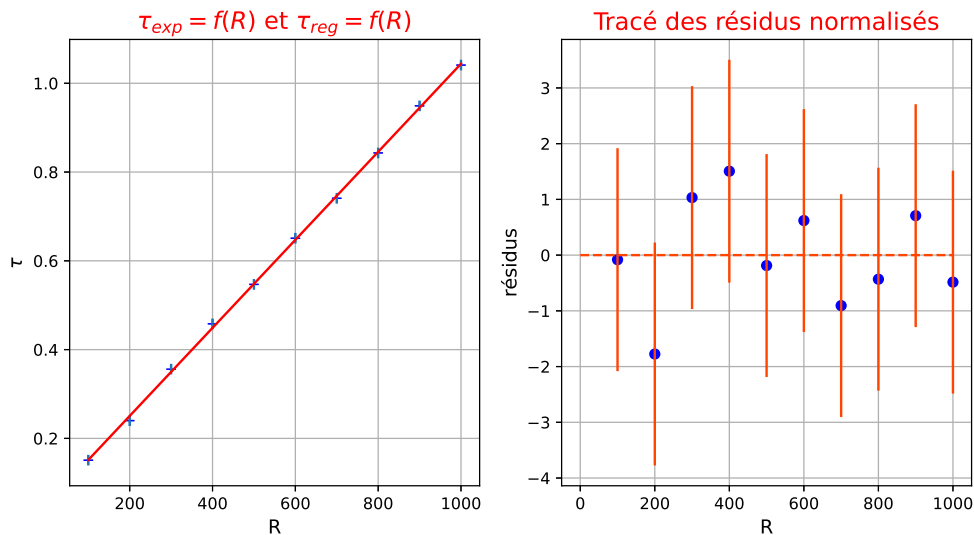
```
1 import numpy as np
2 from matplotlib import pyplot as plt
3 R=np.arange(100,1100,100)
```

```

4 tau=np.array([151e-3, 240e-3, 356e-3, 458e-3, 547e-3, 651e-3, 741e-3, 843e-3, 949e
    -3, 1041e-3])
5 utau=np.array([6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3])
6 N=len(tau)
7 a,b=np.polyfit(R,tau,1)
8 taufit=a*R+b
9 residus=(tau-taufit)/utau #calcul du tableau des résidus normalisés
10 #####
11 ##### Représentations graphiques #####
12 #####
13 plt.subplot(1,2,1) #tracé des points et du modèle affine
14 plt.plot(R,tau,color="blue",marker='+',linestyle='u')
15 plt.errorbar(R,tau,2*utau,linestyle='u')
16 plt.plot(R,taufit,c="red")
17 plt.grid()
18 plt.xlabel("R",fontsize=12)
19 plt.ylabel(r"$\tau$",fontsize=12)
20 plt.title(r"$\tau_{exp}=f(R)$ et $\tau_{reg}=f(R)$",fontsize=15,color="red")
21
22 plt.subplot(1,2,2) #tracé des résidus normalisés
23 plt.plot(R,residus,color="blue",marker='o',linestyle='u')
24 plt.errorbar(R,residus,2,color="orangered",linestyle='u')
25 plt.plot([0,1000],[0,0],linestyle="—",color="orangered")
26 plt.grid()
27 plt.xlabel("R",fontsize=12)
28 plt.ylabel(r"résidus",fontsize=12)
29 plt.title(r"Tracé des résidus normalisés",fontsize=15,color="red")
30
31 print("Valeur de la pente : a={0}F".format(a))
32 print("Valeur de l'ordonnée à l'origine : b={0}us".format(b))

```

On obtient le tracé suivant :



et les valeurs :

In [1]:

Valeur de la pente: $a=0.0009915757575757575$ F

Valeur de l'ordonnée à l'origine: $b=0.052333333333333302$ s)

On constate par le tracé des résidus normalisés que toutes les valeurs respectent le critère sur le z-score.

INCERTITUDE PROPAGÉE SUR a ET b :

dans la mesure où les valeurs (y_1, y_2, \dots, y_N) sont entachées d'une incertitude $u(Y)$, il en va de même pour les grandeurs a et b qui héritent donc d'une incertitude propagée comme toute grandeur calculée à partir de résultats expérimentaux.

⇒ on évaluera donc ces incertitudes propagées par méthode de Monte-Carlo.

On dispose du jeu de N valeurs $(\tau_1, \tau_2, \dots, \tau_N)$ et les incertitudes associées $(u(\tau_1), u(\tau_2), \dots, u(\tau_N))$.

On va réaliser $nbsim$ simulations de l'expérience. Pour chaque simulation, on va faire N tirages aléatoires de τ , chaque tirage $i = 1..N$ étant réalisé par exemple selon une loi normale centrée sur τ_i avec l'écart type $\sigma_i = u_i(\tau)$; on calculera pour chaque tirage de N valeurs les paramètres (a, b) de la droite de régression linéaire; on disposera donc de $nbsim$ valeurs de a et de b sur lesquelles on calculera l'écart-type, estimateur retenu de l'incertitude type.

Le code est le suivant :

Listing I.6 – Détermination de l'incertitude type sur a et b

```

1 import numpy as np
2 from numpy import random as rd
3
4 R=np.arange(100,1100,100)
5 tau=np.array([151e-3, 240e-3, 356e-3, 458e-3, 547e-3, 651e-3, 741e-3, 843e-3, 949e
6             -3, 1041e-3])
7 utau=np.array([6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3, 6e-3])
8 N=len(tau)
9
10 nbsim=int(1e5)
11 tab_a=np.zeros((nbsim),dtype=float)
12 tab_b=np.zeros((nbsim),dtype=float)
13
14 ##### Lancement des nbsim simulations #####
15 for i in range(nbsim):
16     mtau=rd.normal(tau,utau) # grace aux tableaux numpy on traite les N tirages
17     de la simulation i en une seule commande
18     a,b=np.polyfit(R,mtau,1) # on réalise la régression linéaire sur les
19     données de la simulation i en cours
20     tab_a[i],tab_b[i]=a,b #on stocke les valeurs de a et b pour la simulation i
21     en cours
22
23 a_sim,b_sim=np.mean(tab_a),np.mean(tab_b)
24 u_a,u_b=np.std(tab_a,ddof=1),np.std(tab_b,ddof=1) # on calcule l'incertitude type
25 sur a et b
26 print("Valeur de la pente simulée: a_sim={0}F".format(a_sim))
27 print("L'incertitude type sur a vaut: u(a)={0}F".format(u_a))
28 print("Valeur de l'ordonnée à l'origine b simulée: b_sim={0}s".format(b_sim))

```

```
24 | print(r"L'incertitude type sur b vaut: u(b)={0} s".format(u_b))
```

Qui renvoie :

In [85]:

Valeur de la pente a simulée: $a_{\text{sim}}=0.0009915609235465806 \text{ F}$

L'incertitude type sur a vaut: $u(a)=6.6383662412394576e-06 \text{ F}$

Valeur de l'ordonnée à l'origine b simulée: $b_{\text{sim}}=0.05235383371971771 \text{ s}$

L'incertitude type sur b vaut: $u(b)=0.004108847225046534 \text{ s}$

Les valeurs de a et b obtenues par simulation sont bien conformes à celles dégagées par régression linéaire.

En écrivant correctement le résultat, cela donne donc :

$$C = a = (0,9915 \pm 0,0066) \text{ mF}$$

$$b = (0,0523 \pm 0,0041) \text{ s}$$

En ajoutant le code suivant, on trace l'histogramme du tirage des valeurs d'indice :

Listing I.7 –

```
1 | plt.hist(tabn, bins=100)
2 | plt.plot([n_emp, n_emp], [0, 20], c="red")
3 | plt.text(n_emp, 25, r"$n_{\text{emp}}$", c="r")
4 | plt.xlabel(r'n')
5 | plt.ylabel(r'Fréquence')
6 | plt.show()
```

