

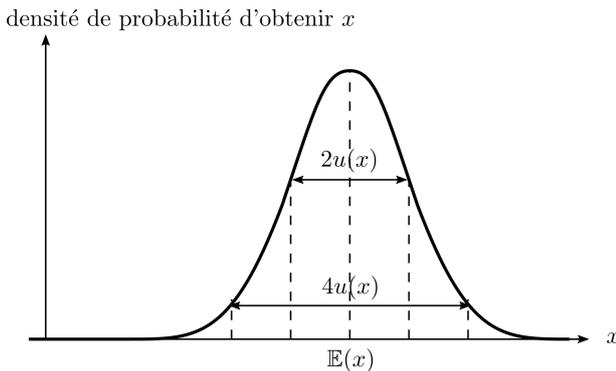
Mesures et incertitudes

1 Variabilité de la mesure et incertitude-type

Lorsqu'on mesure une grandeur physique plusieurs fois dans les mêmes conditions, on n'obtient pas toujours la même valeur. Cette variabilité peut être due à l'opérateur (la personne qui mesure), à l'environnement, aux instruments ou à la méthode de mesure. Pour rendre compte de la **variabilité de la mesure**, on modélise la grandeur d'intérêt x par une **variable aléatoire**. La mesure de x consiste alors à estimer son espérance $\mathbb{E}(x)$ et son écart-type $u(x)$, appelé **incertitude-type**.

En général, la distribution des valeurs mesurées est inconnue. On pourra la modéliser selon les cas par une loi normale ou une loi uniforme.

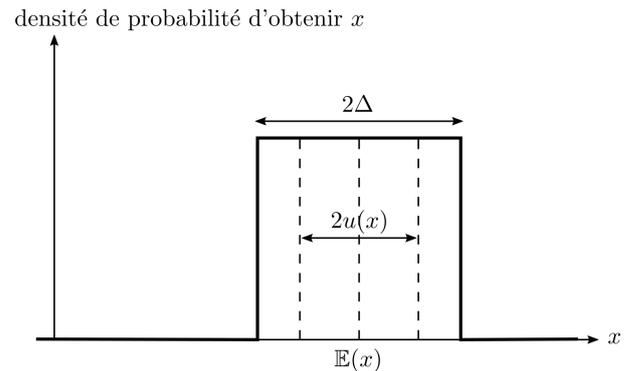
Loi normale



Pour une distribution suivant une loi normale (gaussienne), l'intervalle $[\mathbb{E}(x) - u(x), \mathbb{E}(x) + u(x)]$ est l'**intervalle de confiance à 68%**, c'est-à-dire que la probabilité que la valeur mesurée soit dans cet intervalle est de 68%.

L'intervalle $[\mathbb{E}(x) - 2u(x), \mathbb{E}(x) + 2u(x)]$ est l'intervalle de confiance à 95%.

Loi uniforme



Pour une distribution suivant une loi uniforme de demi-étendue Δ , l'incertitude-type vaut

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

2 Écriture du résultat d'une mesure

Le résultat d'une mesure n'est donc pas une valeur unique, mais un ensemble de valeurs raisonnablement attribuables à la grandeur x , donné sous la forme

- d'une estimation de l'espérance $\mathbb{E}(x)$: c'est la valeur mesurée x ,¹
- **ET** d'une estimation de l'incertitude-type $u(x)$.

On choisit de garder deux chiffres significatifs sur l'incertitude-type $u(x)$. La position du second chiffre significatif de $u(x)$ donne la **position du dernier chiffre significatif à garder sur la valeur mesurée**.

Par exemple, si on a mesuré $L = 1,2$ cm avec une règle graduée au millimètre. On modélise la distribution des mesures par une loi uniforme de demi-étendue $\Delta = 0,5$ mm. L'incertitude-type vaut alors $u(L) = \frac{\Delta}{\sqrt{3}} = 0,29$ cm (avec 2 chiffres significatifs). Le dernier chiffre à garder sur la valeur mesurée est donc le chiffre des millièmes de centimètre. Ainsi, le résultat de cette mesure s'écrit :

$$L = 1,200 \text{ cm avec } u(L) = 0,029 \text{ cm}$$

1. En pratique, on note la valeur mesurée x , comme le mesurande (la grandeur que l'on souhaite mesurer). Cependant, ce ne sont pas les mêmes objets mathématiques : comme son nom l'indique la valeur mesurée est une valeur (avec une unité), tandis que le mesurande est une variable aléatoire.

3 Évaluation de l'incertitude-type

3.1 Évaluation de type A

Lorsqu'on dispose d'une série de valeurs pour une mesure, on peut estimer l'incertitude par une approche statistique. Dans ce cas, on parle d'évaluation de type A. Pour une série de n valeurs mesurées (x_1, x_2, \dots, x_n) obtenues dans les mêmes conditions, la meilleure estimation de l'espérance $\mathbb{E}(x)$ est la moyenne de la série :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

La meilleure estimation de l'incertitude-type sur la moyenne \bar{x} est donnée par :

$$u(\bar{x}) \simeq \frac{\sigma_{n-1}}{\sqrt{n}}$$

où σ_{n-1} est l'écart-type expérimental² de la série, défini par

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

```
import numpy as np
x=np.array([x1,x2,...])
xmoy=np.mean(x)
ux=np.std(x,ddof=1)/np.sqrt(len(x))
# Le diviseur utilisé dans l'écart-type est n-ddof ; par défaut, ddof=0.
print('x=',xmoy,'u(x)=' ,ux)
```

3.2 Évaluation de type B

Lorsque l'incertitude est estimée sur une seule valeur (à partir d'une donnée constructeur, d'une graduation, d'un intervalle estimé par l'opérateur...), on parle d'évaluation de type B.

En général, on n'a aucune information sur la distribution des mesures. Deux situations se présentent souvent :

- L'opérateur (celui qui mesure) peut estimer un intervalle $[x - \Delta, x + \Delta]$ au sein duquel il est « raisonnablement certain » que le résultat de la mesure se trouve. *Par exemple, pour un appareil de mesure analogique (appareil à cadran, lecture sur une règle), Δ peut être pris égal à une demi-graduation.*
- Le constructeur de l'appareil de mesure fournit une indication Δ , sans autre information. *Par exemple, une pipette jaugée possède souvent une indication du type « $\pm 0,04$ mL ».*

Dans ces deux cas, on suppose que la distribution des mesures suit une loi uniforme de demi-étendue Δ . L'incertitude-type est alors donnée par

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

2. L'écart-type expérimental, noté s ou σ_{n-1} , n'est pas tout à fait l'écart-type σ de la série des n valeurs, défini par

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

En effet, on peut montrer mathématiquement, que σ_{n-1} donne une meilleure estimation que σ , de l'incertitude-type $u(x)$. Une interprétation est que, la moyenne \bar{x} étant une estimation de l'espérance $\mathbb{E}(x)$, on dispose en fait d'une donnée de moins que si on calculait l'écart-type directement par rapport à $\mathbb{E}(x)$, d'où la division par $n - 1$.

4 Composition des incertitudes

Lorsqu'une grandeur x s'exprime en fonction d'autres grandeurs y, z , etc, dont les incertitudes-type $u(y), u(z)$, etc sont connues, on peut calculer l'incertitude-type composée $u(x)$.

Attention, avant de se lancer dans un calcul d'incertitude-type composée, il est préférable de comparer entre elles les incertitudes-types relatives $\frac{u(y)}{y}, \frac{u(z)}{z}$, etc des différentes grandeurs d'entrée, et de ne considérer que celles qui sont prépondérantes. Bien souvent, une des grandeurs possède une incertitude relative très supérieure aux autres, ce qui permet d'éviter un calcul fastidieux !

4.1 Calcul analytique

- Somme ou différence : $\boxed{\text{si } x = y + z \text{ ou } x = y - z, \text{ alors } u(x)^2 = u(y)^2 + u(z)^2}$
- Produit ou quotient³ : $\boxed{\text{si } x = yz \text{ ou } x = \frac{y}{z}, \text{ alors } \left[\frac{u(x)}{x}\right]^2 = \left[\frac{u(y)}{y}\right]^2 + \left[\frac{u(z)}{z}\right]^2}$

4.2 Calcul par simulation Monte-Carlo

Lorsque l'expression $x = f(y, z, \dots)$ est plus compliquée, on calcule $u(x)$ numériquement, en simulant un processus aléatoire (simulation dite *Monte-Carlo*).

1. Pour chaque grandeur d'entrée y, z , etc, on simule un grand nombre N de mesures aléatoires.
 - Si on n'a aucune autre information qu'une limite basse $y - \Delta_y$ et une limite haute $y + \Delta_y$ pour la grandeur y , alors on suppose que la distribution suit une loi uniforme entre ces deux bornes. La fonction `uniform(a,b,N)` du module `numpy.random` renvoie un tableau de N valeurs suivant une loi uniforme sur $[a, b]$.
 - Si on connaît la valeur mesurée y et l'incertitude-type associée $u(y)$, mais qu'on ne connaît pas la distribution sous-jacente, alors on suppose que la distribution suit une loi normale. La fonction `normal(y,uy,N)` du module `numpy.random` renvoie un tableau de N valeurs suivant une loi normale d'espérance y et d'écart-type $u(y)$.
2. On calcule ensuite $x_i = f(y_i, z_i, \dots)$ pour chaque tirage $i \in \llbracket 0, N - 1 \rrbracket$.
3. Finalement, l'incertitude-type composée $u(x)$ est donnée par l'écart-type de la série des x_i . N étant très grand, on peut calculer indifféremment l'écart-type σ , ou l'écart-type expérimental σ_{n-1} .

```
import numpy as np
import numpy.random as rd

y=...
uy=... # ou Deltay=...
z=...
uz=... # ou Deltaz=...

N=10**6
yMC=rd.normal(y,uy,N) # dans le cas où on dispose de l'incertitude-type sur y
# ou yMC=rd.uniform(y-Deltay,y+Deltay,N) dans le cas où on dispose de la demi-étendue
zMC=rd.normal(z,uz,N) # ou zMC=rd.uniform(z-Deltaz,z+Deltaz,N)
xMC=f(yMC,zMC) # f à définir...

x=f(y,z) # valeur mesurée (inutile de calculer la moyenne de xMC !)
ux=np.std(xMC) # incertitude-type
print('x=' , x, 'u(x)=' , ux)
```

-
3. De cette formule, on peut également déduire
 - Multiplication par une constante : si $x = \lambda y$, alors $u(x) = |\lambda|u(y)$
 - Puissance : si $x = y^\alpha$, alors $\left[\frac{u(x)}{x}\right]^2 = |\alpha| \left[\frac{u(y)}{y}\right]^2$

5 Comparaison de deux valeurs

Pour comparer deux valeurs x_1 et x_2 dont les incertitudes-type $u(x_1)$ et $u(x_2)$ sont connues, on peut comparer l'écart entre les deux valeurs $|x_1 - x_2|$, à l'incertitude-type sur cet écart $\sqrt{u(x_1)^2 + u(x_2)^2}$.

On définit ainsi l'**écart normalisé** (ou z-score) :

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

- Si $E_N \leq 2$ on considérera que les résultats des deux mesures sont compatibles.
- Si $E_N > 2$ on considérera que les résultats des deux mesures sont incompatibles.⁴ Il convient alors d'analyser les causes de cette incompatibilité (approximation dans la modélisation, biais dans les mesures, sous-estimation des incertitudes, etc).

6 Régression linéaire

6.1 Validation d'un modèle théorique

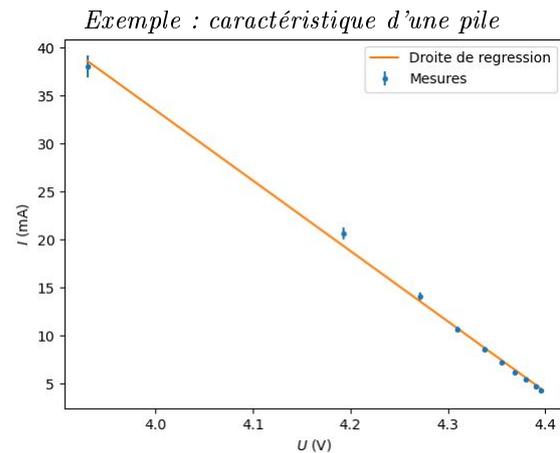
On cherche à valider une loi théorique affine, de la forme $y = ax + b$.⁵ Pour cela, on réalise une série de mesures de couples (x_i, y_i) . Pour chaque mesure, on estime l'incertitude-type $u(y_i)$, par une évaluation de type A ou B. Les valeurs des paramètres a et b peuvent alors être estimées par une régression linéaire, avec la fonction `polyfit` du module `numpy`. On peut ensuite tracer le graphe $y(x)$, sur lequel on superpose les points de mesures avec leurs barres d'incertitude et la droite de régression.

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

x=np.array([x1,x2,...])
y=np.array([y1,y2,...])
uy=np.array([uy1,uy2,...])

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

plt.close()
plt.errorbar(x,y,yerr=uy,fmt='.') # mesures
plt.plot(x,a*x+b) # droite (modèle)
plt.show()
```



On appelle résidus, les écarts des points de mesure à la droite de régression : $y_i - (ax_i + b)$.

On considère que les mesures et le modèle sont compatibles, si chaque résidu est inférieur à (ou de l'ordre de) $2u(y_i)$, en valeur absolue.

Lorsque certaines barres d'incertitude sont trop petites, pour qu'on puisse voir à l'œil si les mesures et le modèle sont compatibles, on peut représenter le graphe des résidus normalisés

$$R_{Ni} = \frac{y_i - (ax_i + b)}{u(y_i)}$$

en fonction de x .

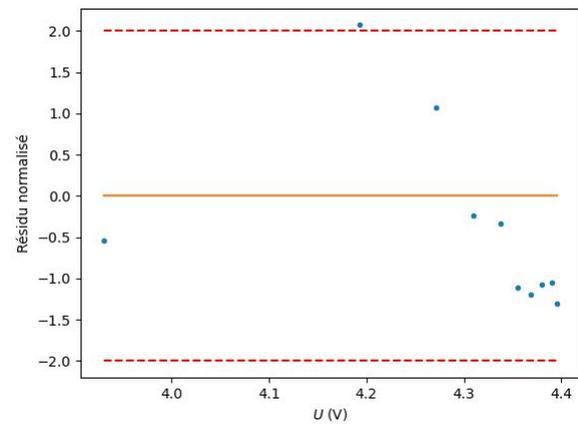
4. Le seuil de 2 est un choix arbitraire. En réalité, si $E_N \gtrsim 2$, on ne peut pas définitivement conclure ! En effet, même pour des mesures compatibles, la probabilité que $E_N > 2$ est d'environ 5%. Ainsi, si on effectue une série de 20 mesures dans les mêmes conditions, il devient probable qu'une d'entre elles ait un écart normalisé au résultat attendu supérieur à 2.

5. Pour choisir entre x et y , la grandeur en abscisse et la grandeur en ordonnée, il faut comparer les incertitudes-types relatives $\frac{u(x_i)}{x_i}$ et $\frac{u(y_i)}{y_i}$ sur ces grandeurs. On choisit comme ordonnée la grandeur y dont les incertitudes-types relatives sont les plus grandes. On néglige les incertitudes-types $u(x_i)$ sur la grandeur x en abscisse.

```

RN=(y-a*x-b)/uy
plt.close()
plt.plot(x,RN,'.i')
plt.show()

```



On considère que les mesures et le modèle sont compatibles, si chaque résidu normalisé est inférieur à (ou de l'ordre de) 2, en valeur absolue.

6.2 Évaluation des incertitudes-type sur a et b par simulation Monte-Carlo

Lorsqu'on fait une régression linéaire entre deux grandeurs x et y , pour lesquelles on dispose d'une série de n couples de mesures (x_i, y_i) , on peut estimer les incertitudes-types sur le coefficient directeur a et l'ordonnée à l'origine b en simulant un processus aléatoire (simulation *Monte-Carlo*).

1. On rentre les valeurs mesurées des x_i et des y_i dans deux listes. On rentre dans une troisième liste les incertitudes-types $u(y_i)$, ou les demi-étendues Δ_i des distributions des valeurs de chaque y_i , selon ce dont on dispose.
2. On simule ensuite un grand nombre N de séries de n mesures aléatoires (y_1, y_2, \dots, y_n) , en faisant l'hypothèse d'une loi normale ou uniforme, selon que l'on dispose des incertitudes-types $u(y_i)$ ou des demi-étendues Δ_i des distributions des y_i . Pour chaque série de n mesures, une régression linéaire, avec la fonction `polyfit` du module `numpy`, permet d'obtenir un coefficient directeur a et une ordonnée à l'origine b , que l'on stocke dans deux listes. On obtient ainsi deux listes de N valeurs.
3. Finalement, les écarts-types des listes de N valeurs de a et de b donnent des estimations des incertitudes-types $u(a)$ et $u(b)$.

```

import numpy as np
import numpy.random as rd

x=np.array([x1,x2,...]) # liste des n mesures de x
y=np.array([y1,y2,...]) # liste des n mesures de y
uy=np.array([uy1,uy2,...]) # dans le cas où on dispose des incertitudes-types sur les yi
# ou Delta=np.array([Delta1,Delta2,...]) dans le cas où on dispose des demi-étendues
N=10**4

a,b=np.polyfit(x,y,1) # coefficient directeur, ordonnée à l'origine

aMC=[] # liste des coefficients directeurs simulés
bMC=[] # liste des ordonnées à l'origine simulées
for j in range(N) :
    yMC=rd.normal(y,uy) # ou yMC=rd.uniform(y-Delta,y+Delta)
    p=np.polyfit(x,yMC,1)
    aMC.append(p[0])
    bMC.append(p[1])
ua=np.std(aMC)
ub=np.std(bMC)

print('a=',a,'u(a)=' ,ua)
print('b=',b,'u(b)=' ,ub)

```