

NOTIONS de BASE
sur les INCERTITUDES et le
TRAITEMENT des DONNEES EXPERIMENTALES
en
PHYSIQUE, CHIMIE, BIOLOGIE

POSITION du PROBLEME

Toutes les sciences, et particulièrement les sciences pour l'ingénieur, reposent sur la confrontation entre des résultats d'*expériences* supposées représentatives et de *théories* censées décrire objectivement la réalité du monde. Les deux démarches, expérimentale ou théorique, sont naturellement accompagnées de la notion de doute, doute qui est un concept fondamental pour un (futur) ingénieur.

Le but de ce texte est de présenter les notions liées aux jugements sur la qualité des déterminations expérimentales. En effet, il n'est pas possible de mesurer une grandeur physique, chimique, biologique... qui soit totalement exempte d'erreur ou d'incertitude. Pourtant, *tout résultat* présente un intérêt à condition qu'il soit exprimé avec ses limites d'incertitude.

Il n'existe malheureusement pas de méthode simple pour déterminer la validité de mesures expérimentales. Ce polycopié donne quelques pistes générales, dans un nombre limité de pages, afin que l'expérimentateur puisse présenter ses résultats avec *bon sens*.

Pour des applications particulières, la consultation d'ouvrages spécialisés tels que ceux proposés en bibliographie est indispensable.

I- VOCABULAIRE

On confond souvent dans le langage scientifique "erreur" et "incertitude" même si on utilise plutôt le terme "erreur" lorsqu'un processus de mesure est mal maîtrisé et "incertitude" lorsque l'évaluation de la fiabilité est immédiate et intuitive.

Par contre le vocabulaire est précis pour un certain nombre de notions que vous devez connaître:

1) Incertitude absolue, Incertitude relative:

L'incertitude absolue Δx est l'erreur maximale que l'on est susceptible de commettre dans l'évaluation de x . L'incertitude absolue s'exprime donc dans les unités de la grandeur mesurée.

Exemple 1 : Les physiciens américains Dumond et Cohen ont proposé au début des années 1950 plusieurs valeurs expérimentales pour la vitesse de la lumière :

- 1948 : $c = (299776 \pm 4) \text{ km.s}^{-1}$
- 1951 : $c = (299790,0 \pm 0,9) \text{ km.s}^{-1}$
- 1953 : $c = (299792,9 \pm 0,8) \text{ km.s}^{-1}$.

On remarquera l'incompatibilité apparente entre ces résultats, particulièrement entre celui de 1948 et ceux de 1951 et 1953. La conclusion est que ces chercheurs ont été optimistes dans l'appréciation des incertitudes...

L'incertitude relative $\Delta x/x$ représente l'importance de l'erreur par rapport à la grandeur mesurée. L'incertitude relative n'a pas d'unités et s'exprime en général en % ($100\Delta x/x$).

Exemple 2: une balance d'analyse de laboratoire permet de peser typiquement à $\pm 0,1$ mg près.

Si la pesée est de 10 mg l'incertitude absolue est $\pm 0,1$ mg. L'incertitude relative est 1%.

Si la pesée est de 1000 mg l'incertitude absolue est toujours $\pm 0,1$ mg. L'incertitude relative est 0,01%. Mais on peut choisir de peser à 10 mg près si on se contente d'une incertitude relative de 1%.

Exemple 3: La mesure de la vitesse de la lumière exprimée par $c = (299792,9 \pm 0,8) \text{ km.s}^{-1}$ correspond à une incertitude relative $\Delta c/c = 3.10^{-6}$.

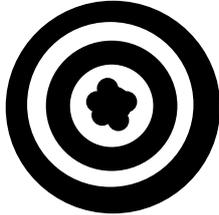
L'incertitude relative est une manière commode de chiffrer la *précision* d'une mesure (notion définie ci dessous). Dans le langage scientifique, une mesure de fréquence est dite *très précise* car mesurée avec une incertitude relative meilleure que 0,0001 %. Par contre la mesure d'un paramètre physiologique ne peut être difficilement apprécié à mieux que quelques dizaines de % du fait de la variabilité rencontrée chez les êtres vivants.

2) précision et exactitude:

Dans le langage habituel, ces deux derniers termes sont souvent confondus à tort.

La *précision* traduit le degré de proximité (ou reproductibilité) que l'on observe entre différentes mesures qui ont été obtenues par la même méthode.

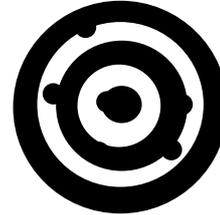
L'*exactitude* exprime la proximité entre un résultat et sa valeur réelle ou présumée telle. Cette notion nécessite la comparaison entre différentes méthodes.



Bonne précision
Bonne exactitude



Bonne précision
Mauvaise exactitude



Bonne exactitude
Mauvaise précision

Exemple 4: autre illustration du cas « bonne précision-mauvaise exactitude » ci dessus. Une nouvelle méthode est en cours de mise au point pour le dosage d'un composé actif de médicament par chromatographie (HPLC). La méthode est trouvée précise car la dispersion des mesures est de l'ordre de 1% mais son exactitude est mauvaise car la valeur moyenne est mesurée systématiquement 10% au dessous de la valeur moyenne obtenue par la méthode de référence (par exemple un marquage radioactif).

L'idéal est évidemment une mesure à la fois précise et exacte. Un moindre mal est une mesure exacte et imprécise.

3) Les types d'erreurs dans les données expérimentales:

On peut distinguer trois types principaux d'erreurs:

a) *Les erreurs aléatoires* (ou fortuites) affectent la précision des mesures. Elles sont étudiées soit en effectuant des calculs d'incertitudes soit en comparant statistiquement les résultats d'expériences soigneusement répétées.

b) *Les erreurs grossières* se produisent rarement et sont facilement identifiées car elles conduisent à des valeurs excentriques qui diffèrent significativement des résultats dans une série de mesures. Elles sont naturellement éliminées avant interprétation des résultats.

c) Il est beaucoup plus délicat d'identifier *les erreurs systématiques* qui affectent l'exactitude. Par exemple une erreur systématique s'est manifestement glissée dans la méthode chromatographique de l'exemple 4. Elle pourrait être due, par exemple, à un mauvais rendement d'extraction du composé actif dans le médicament.

Il est difficile d'être exhaustif dans la description des erreurs systématiques. Les principales sont:

- *Les erreurs de méthode.*

Exemple 5: en volumétrie, le virage d'un indicateur coloré peut se produire avant ou après l'équivalence théorique.

- *Les erreurs instrumentales.*

Exemple 6: une pipette jaugée insuffisamment nettoyée ne délivre pas le volume nominal. Un convertisseur analogique-digital peut être mal étalonné.

- *Les erreurs personnelles.*

Exemple 7: lors de l'utilisation d'une fiole jaugée, le ménisque peut être systématiquement ajusté trop haut.

Les erreurs systématiques sont les plus difficiles à détecter et nécessitent une vigilance constante dans les laboratoires. Certaines procédures dites "bonne pratique de laboratoire" permettent de donner un cadre réglementaire à cette vigilance.

Dans la suite de ce texte seul le traitement des erreurs aléatoires sera détaillé.

4) Le "coût" de la précision:

L'obtention de la meilleure précision possible n'est pas une fin en soi. L'important quand on entreprend une série de mesures, est de savoir à quoi elles vont servir, d'adapter l'appareillage à la précision souhaitée, au moindre coût et le plus rapidement possible (voir l'exemple 11).

II- Traitement non statistique des incertitudes (erreurs aléatoires)

La méthode la plus rigoureuse d'évaluation des erreurs aléatoires est toujours la méthode statistique (voir III), mais elle exige de répéter un nombre de fois significatif l'expérience et ce n'est pas toujours possible.

Une manière simple d'appréhender l'incertitude sur un résultat est d'utiliser la combinaison des incertitudes de chaque étape, c'est ce qu'on appelle un calcul d'incertitude. Ce calcul n'est lui aussi pas toujours possible car l'identification d'une "relation expérimentale" qui lie toutes les grandeurs mesurées n'est pas toujours évidente. Par exemple, lors d'un protocole d'expérimentation sur des animaux de laboratoire, seule une méthode statistique permet d'évaluer la précision de la mesure.

1) le calcul d'incertitude:

Il se décompose en 3 étapes:

a) identification de la *relation expérimentale* qui doit expliciter toutes les grandeurs utilisées:

Exemple 8: une solution $C_1 = 5,5 \cdot 10^{-3} \text{ mol l}^{-1}$ est préparée à partir d'une solution mère $C_2 = 10^{-2} \text{ mol l}^{-1}$ par une double prise d'essai de $v_1 = 5 \text{ cm}^3$ et $v_2 = 0,5 \text{ cm}^3$ complétée à $V = 10 \text{ cm}^3$.

La relation expérimentale qui explicite le calcul de C_1 permettra de calculer l'incertitude sur C_1 :

$$C_1 = \frac{v_1 + v_2}{V} C_2$$

b) identification de chacune des incertitudes intermédiaires:

Les incertitudes dépendent du matériel utilisé.

Dans l'exemple 8 précédent la solution mère C_2 est une solution commerciale étalon (garantie d'une précision de 0,2%), soit:

$$C_2 = 10^{-2} \text{ mol l}^{-1} \quad \Delta C_2 = 0,002 \cdot 10^{-2} \text{ mol l}^{-1} \quad \Delta C_2 / C_2 = 0,2\%$$

v_1 est prélevé par une pipette jaugée de classe A (incertitude relative = 0,2 %) soit:

$$v_1 = 5 \text{ cm}^3 \quad \Delta v_1 = 0,01 \text{ cm}^3 \quad \Delta v_1 / v_1 = 0,2\%$$

v_2 est prélevé par une pipette graduée de 1 cm^3 , de classe A (incertitude absolue = 1% du volume total) soit:

$$v_2 = 0,5 \text{ cm}^3 \quad \Delta v_2 = 1\% \text{ de } 1 \text{ cm}^3 \text{ soit } = 0,01 \text{ cm}^3 \quad \Delta v_2 / v_2 = 2\%$$

V est obtenu par ajustage d'une fiole jaugée de classe A (incertitude relative = 0,2 %) soit:

$$V = 10 \text{ cm}^3 \quad \Delta V = 0,02 \text{ cm}^3 \quad \Delta V / V = 0,2\%$$

c) calcul de l'incertitude selon quelques règles simples:

La relation expérimentale est différenciée par rapport à chacune des grandeurs, considérées comme indépendantes. Dans le cas de produits ou de quotients il est rapide d'effectuer cette différenciation en passant par les logarithmes.

Pour l'exemple 8 en cours, on trouve:

$$\frac{dC_1}{C_1} = \frac{dv_1}{(v_1 + v_2)} + \frac{dv_2}{(v_1 + v_2)} + \frac{dC_2}{C_2} - \frac{dV}{V}$$

Si nécessaire il faut regrouper les termes correspondants à la même variable.

Le passage aux incertitudes correspond au passage à la plus grande valeur possible en valeur absolue de tous les coefficients multiplicatifs, soit ici:

$$\frac{\Delta C_1}{C_1} = \Delta v_1 \left| \frac{1}{(v_1 + v_2)} \right| + \Delta v_2 \left| \frac{1}{(v_1 + v_2)} \right| + \Delta C_2 \left| \frac{1}{C_2} \right| + \Delta V \left| - \frac{1}{V} \right|$$

ou encore si les quantités sont positives:

$$\frac{\Delta C_1}{C_1} = \frac{\Delta v_1}{(v_1 + v_2)} + \frac{\Delta v_2}{(v_1 + v_2)} + \frac{\Delta C_2}{C_2} + \frac{\Delta V}{V}$$

L'application numérique donne:

$$\Delta C_1 / C_1 = 0,01/5,5 + 0,01/5,5 + 0,2/100 + 0,2/100 = 0,8 \% \text{ (on arrondit toujours à la valeur supérieure).}$$

$$\Delta C_1 = 0,044 \cdot 10^{-3} \text{ mol l}^{-1}.$$

soit en conclusion: $C_1 = 5,50 \cdot 10^{-3} \pm 0,05 \cdot 10^{-3} \text{ mol l}^{-1}$ ou encore $C_1 = (5,50 \pm 0,05) \cdot 10^{-3} \text{ mol l}^{-1}$.

Ce type de calcul est facilité par des règles simples qui se démontrent aisément à partir de ce qui précède:

L'incertitude absolue sur une somme algébrique de mesures est la somme des incertitudes absolues sur chacun des termes.

Exemple 9: pour la relation expérimentale $A = a-b+2c$ on trouve $\Delta A = \Delta a + \Delta b + 2\Delta c$

L'incertitude relative sur le produit ou(et) le quotient de mesures indépendantes est la somme des incertitudes relatives (affectés des coefficients nécessaires).

Exemple 10: pour la relation expérimentale: $X = (2u + 1)^3 \sqrt{\frac{y}{z^2}}$ le calcul par les logarithmes donne:

$$\frac{\Delta X}{X} = \frac{2\Delta u}{2u+1} + \frac{1}{3} \left(\frac{\Delta y}{y} \right) + \frac{2}{3} \left(\frac{\Delta z}{z} \right)$$

Exemple complet N°11: détermination de la résistivité électrique ρ du cuivre.

• Identification de la loi expérimentale : un fil électrique de diamètre d , de longueur l et de résistance électrique R est réalisé en cuivre dont la conductivité électrique est donnée par : $\rho = \frac{\pi}{4} R \frac{d^2}{l}$.

• Identification des incertitudes intermédiaires : les résultats des mesures nous donnent $d = (0,30 \pm 0,01) \text{ mm}$, $l = (2 \pm 0,001) \text{ m}$ et $R = (0,4562 \pm 0,0002) \Omega$.

• Calcul de l'incertitude sur la résistivité électrique. Le théorème des incertitudes relatives donne :

$$\frac{\Delta \rho}{\rho} = 2 \frac{\Delta d}{d} + \frac{\Delta l}{l} + \frac{\Delta R}{R}$$

Les trois termes intervenant dans le calcul sont d'importance inégale. En effet nous avons : $2\Delta d/d = 7 \cdot 10^{-2}$, $\Delta l/l = 5 \cdot 10^{-4}$ et $\Delta R/R = 3 \cdot 10^{-4}$. Nous pouvons donc négliger $\Delta l/l$ et $\Delta R/R$ devant $\Delta d/d$. Donc : $\Delta \rho/\rho = 2\Delta d/d = 7 \cdot 10^{-2}$.

• Présentation finale du résultat : le calcul de ρ peut alors être effectué en prenant $\pi = 3,14$. Nous obtenons : $\rho = (1,6 \pm 0,1) \cdot 10^{-8} \Omega \cdot \text{m}$.

Pour améliorer la mesure de ρ , il faut diminuer si possible l'incertitude sur la mesure du diamètre du fil. Il est alors inutile de mesurer l et surtout R avec une aussi grande précision.

2) Présentation de résultats:

Lors de la présentation finale d'un résultat il est important d'accorder le nombre de chiffres significatifs à la précision déterminée.

Dans le cas où une incertitude n'est pas explicitement donnée, les scientifiques admettent le niveau du dernier chiffre significatif comme ordre de grandeur de l'incertitude.

Exemple 12: Vous trouvez dans une table de constantes:

$A = 23,0$ unité. Vous interprétez que A est connu à $\pm 0,1$ ou $0,2$ unité près (incertitude relative 0,5 à 1%)

$B = 0,007$ unité. Vous interprétez que B est connu à $\pm 0,001$ unité près (incertitude relative 15%)

$C = 59000$ unité. Vous interprétez que C est connu à ± 1000 unité près (incertitude relative 1,7%).

Ceci ne transmet qu'un ordre de grandeur de l'incertitude mais c'est déjà important.

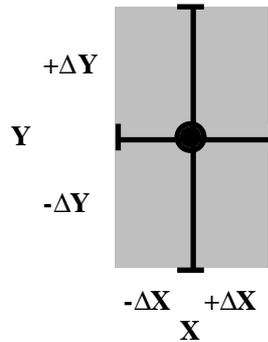
3) Application au dépouillement de résultats graphiques

Il est courant d'étudier graphiquement une propriété en fonction d'un paramètre pour en déduire ou vérifier une loi (linéaire par exemple).

Les rectangles d'incertitudes (ou barres d'erreurs) doivent être portés sur le graphe pour juger de la validité de l'interprétation.

Soit un point expérimental défini par les coordonnées: X affecté de l'incertitude $\pm \Delta X$
Y affecté de l'incertitude $\pm \Delta Y$

Le tracé de ce point sur un graphe correspond au schéma suivant:



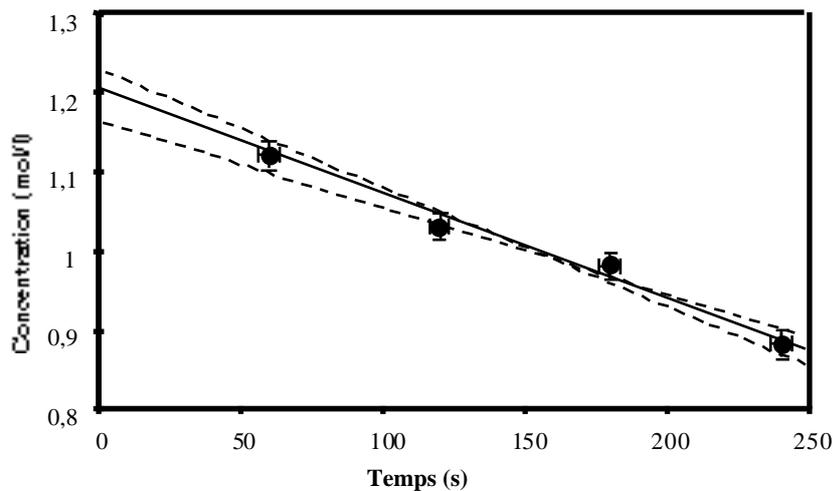
La zone grise correspond à l'aire d'incertitude du point expérimental. Elle peut se réduire à une simple barre si l'une des incertitudes est très faible.

Exemple 13: interprétation de résultats graphiques:

Pour étudier une cinétique chimique dans un réacteur, la concentration d'une espèce est suivie par prélèvements réguliers en fonction du temps. Le dosage aboutit à une mesure de la densité optique (absorbance) de la solution. Compte tenu de toutes les contraintes expérimentales, les valeurs mesurées sont entachées des incertitudes suivantes:

Temps = t (seconde)	60 ± 3	120 ± 3	180 ± 3	240 ± 3
Concentration = C (mol/l)	1,12 $\pm 0,03$	1,03 $\pm 0,03$	0,98 $\pm 0,02$	0,88 $\pm 0,02$

Ce qui donne le graphique:



Pour l'interprétation, on trace manuellement la meilleure droite passant au mieux dans tous les rectangles d'incertitude. On trouve l'équation: $C_1 = 1,20 - 1,31 \times 10^{-3} t$

Puis, de même, les deux droites de pentes extrêmes possibles (en pointillés), respectivement d'équation:

$$C_2 = 1,23 - 1,50 \times 10^{-3} t$$

$$C_3 = 1,16 - 1,06 \times 10^{-3} t$$

On peut donc en déduire:

pour $t = 0$ l'extrapolation de la concentration donne: $C_0 = 1,20 \pm 0,04 \text{ mol l}^{-1}$

La pente de la droite est $1,3 \pm 0,3 \text{ mol l}^{-1}\text{s}^{-1}$.

Note: cette manière de procéder conduit en général à une **surestimation** de l'incertitude sur le résultat final. Elle se justifie si on dispose uniquement de 3 ou 4 points mais dès que le nombre de points est suffisant il vaut mieux utiliser un traitement statistique (régression linéaire).

III- Traitement statistique des erreurs aléatoires

Il n'est pas possible de donner en quelques lignes les principes souvent très élaborés des traitements statistiques dont le but est de permettre un jugement objectif d'un ensemble de résultats même si les incertitudes ne sont pas a priori connues. Après lecture des quelques lignes de présentation qui suivent, la consultation d'un ouvrage spécialisé (voir bibliographie) est indispensable.

1) principes de base:

On appelle *échantillon* de données un nombre fini (N) d'observations expérimentales de la grandeur x. Comme les lois statistiques ont été établies pour une *population* c'est à dire un échantillon de très grande taille, elles doivent être modifiées quand on les applique à un petit échantillon de données.

On appelle *moyenne* de l'échantillon le terme suivant:

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

L'écart type (s) d'un échantillon fini est défini par:

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

Exemple 14:

L'analyse de la teneur en plomb d'un échantillon de sang a donné les résultats suivants:

0,752 0,756 0,752 0,751 0,760 en ppm (partie par million).

Les formules précédentes pour $N = 5$ donnent: $\bar{x} = 0,754 \text{ ppm}$ et $s = 0,0038 \quad 0,004 \text{ ppm}$

On peut approfondir la signification statistique de l'écart type (par exemple pour exprimer la précision) à condition d'y associer une *limite de confiance* (LC). La limite de confiance définit un intervalle de confiance dans lequel la valeur réelle a p % de chance de se trouver.

Les calculs théoriques démontrent que l'écart type doit être majoré par un facteur multiplicatif d'autant plus grand que le nombre de mesures est faible et l'exigence de fiabilité grande:

A une limite de confiance de p % correspond l'intervalle $= \bar{x} \pm \frac{ts}{\sqrt{N}}$

où N, s, \bar{x} , ont les significations précédentes et où t est le t *statistique* ou t *de Student*.

On trouve dans les ouvrages de statistique des tableaux complets donnant t en fonction de N et des limites de confiance.

Nombre de mesures N	Facteur p pour les limites de confiance			
	80%	90%	99%	99,9%
2	3,08	6,31	63,7	637
3	1,89	2,92	9,92	31,6
4	1,64	2,35	5,84	12,9
5	1,53	2,13	4,60	8,60
6	1,48	2,02	4,03	6,86
10	1,38	1,83	3,25	4,78
15	1,34	1,76	2,98	4,14
	1,29	1,64	2,58	3,29

On retrouve dans ce tableau la notion intuitive qu'il faut multiplier les mesures pour augmenter leur fiabilité.

Dans l'exemple 14 en cours, N=5. Si on choisit une limite de confiance de 99% on trouve t = 4,60

$$\text{soit } \frac{ts}{\sqrt{N}} = 4,6 \times 0,0038 / 2,236 = 0,0078 \quad 0,008 \text{ ppm}$$

La conclusion est qu'il y a 99% de chance que la valeur soit dans l'intervalle: $0,754 \pm 0,008$ ppm.

On note : LC à 99% = $0,754 \pm 0,008$ ppm

Dans le cas ou un résultat final fait intervenir plusieurs écarts types sur des grandeurs indépendantes, ces écarts types peuvent être combinés un peu comme dans les calculs d'incertitudes. Néanmoins, les règles sont différentes et plus complexes. Il faut consulter un ouvrage spécialisé.

2) Analyse de régression par la méthode des moindres carrés:

Le tracé de la meilleure droite au milieu de points expérimentaux est un problème courant au laboratoire. Une méthode numérique objective appelée *analyse de régression* permet d'établir cette droite et de calculer les incertitudes associées. On ne considère ici que la méthode de régression la plus simple, la *méthode des moindres carrés*.

- *Hypothèses:*

Soit une relation linéaire $y = mx + b$.

On considère que x est affecté d'incertitude négligeable mais que y est entaché d'erreurs expérimentales aléatoires. Ce cas simple s'applique facilement dans le cas d'une courbe d'étalonnage.

- *Calcul de la droite des moindres carrés:*

L'écart vertical de chaque point à la droite est appelé un *résidu*. La droite calculée par la méthode des moindres carrés est celle qui minimise la somme des carrés des résidus de tous les points (voir ouvrage spécialisé).

En pratique on définit trois grandeurs S_{xx} , S_{yy} , S_{xy} :

$$S_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{N}$$

$$S_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{N}$$

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{\sum x_i \sum y_i}{N}$$

où x_i et y_i sont des paires de données, N le nombre de paires et \bar{x} et \bar{y} sont les valeurs moyennes des variables.

On peut en déduire:

- la pente de la droite:

$$m = S_{xy} / S_{xx}$$

- l'ordonnée à l'origine: $b = \bar{y} - m\bar{x}$

- l'écart type des résidus: $s_r = \sqrt{\frac{S_{yy} - m^2 S_{xx}}{N - 2}}$

- l'écart type de la pente: $s_m = \sqrt{\frac{s_r^2}{S_{xx}}}$

- l'écart type de l'ordonnée à l'origine: $s_b = s_r \sqrt{\frac{\sum x_i^2}{N \sum x_i^2 - (\sum x_i)^2}}$

Ces calculs sont facilement accessibles sur calculette. par ailleurs il existe de nombreux logiciels sur micro-ordinateur, tableurs ou graphiques, intégrant ces calculs avec le tracé automatique des points et de la droite ce qui permet un contrôle visuel du résultat.

Les écarts types obtenus sont analogues à l'écart type de données unidimensionnelles. On peut en déduire d'une manière similaire des limites de confiance comme dans 1).

-Le coefficient de corrélation: r ou coefficient de Pearson permet d'établir le niveau de validité de la relation linéaire entre les deux variables x et y:

$$r = \frac{N \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{[N \sum x_i^2 - (\sum x_i)^2][N \sum y_i^2 - (\sum y_i)^2]}}$$

Plus la valeur de r est proche de 1 plus grande est la probabilité d'une corrélation de linéarité entre les deux variables x et y. Ainsi un jugement sur la valeur de r peut être porté à partir de tables de valeurs critiques de r (voir bibliographie).

Néanmoins, le seul calcul de r ne peut suffire, il faut absolument vérifier graphiquement comment la droite passe au milieu des points. Ce contrôle est d'autant plus important que le nombre de points est faible.

Exemple 15:

La courbe d'étalonnage d'une protéine séparée par chromatographie (HPLC) est obtenue par intégration du pic pour différentes solutions étalons de cette protéine:

x = Concentration (mg/l)	0,352	0,803	1,08	1,38	1,75
y = Aire du pic (unité arbitraire)	1,09	1,78	2,60	3,03	4,01

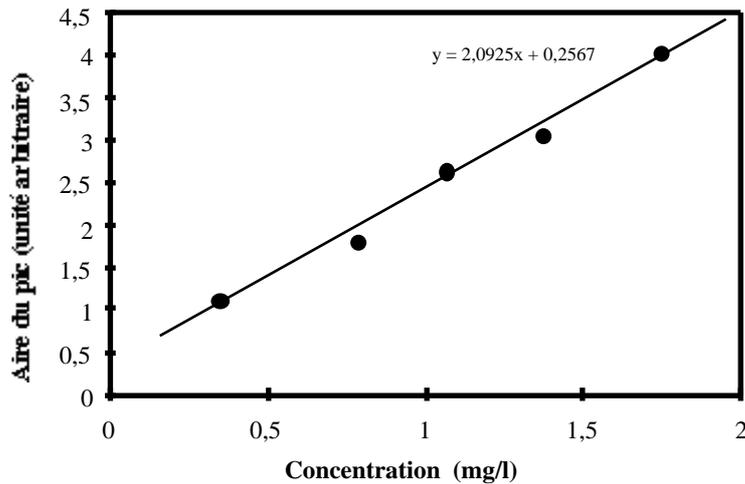
L'application des relations précédentes donne numériquement:

$$S_{xx} = 6,90201 - (5,365)^2/5 = 1,14537$$

$$S_{yy} = 36,3775 - (12,51)^2/5 = 5,07748$$

$$S_{xy} = 15,81992 - 5,365 \times 12,51/5 = 2,39669$$

L'équation de la droite des moindres carrés est donc: $y = 2,09x + 0,26$ ce qui donne graphiquement (par exemple en utilisant EXCEL ou votre calculatrice):



avec les écarts types: $s_r = 0,14$ $s_m = 0,13$ $s_b = 0,16$

Si on veut exprimer, par exemple la pente, avec une limite de confiance donnée, par exemple à 90% on trouve dans le tableau du III,1: $t = 2,13$ soit $ts_m/\sqrt{N} = 0,12$.

soit: LC à 90% pour $m = 2,09 \pm 0,12$

Par ailleurs le coefficient de corrélation trouvé est $r = 0,994$, ce qui permettrait de juger (voir bibliographie) que le résultat trouvé est un résultat fiable.

Il existe des logiciels plus élaborés (par exemple Minim 3) qui permettent d'attribuer un poids différent (donc une incertitude différente) à chaque point expérimental. Les calculs de régression ou d'identification de paramètres sur ces logiciels permettent d'affiner la discussion dans les cas difficiles.

CONCLUSION

Il existe donc plusieurs approches pour évaluer la validité de résultats expérimentaux.

- Au minimum, un résultat doit être exprimé avec une incertitude calculée à l'aide des calculs d'incertitudes simples. Mais l'inconvénient est qu'il n'est pas toujours possible d'identifier la relation expérimentale indispensable pour effectuer ce calcul.
- L'utilisation de méthodes statistiques permet d'exprimer plus rigoureusement les limites de fiabilité du résultat. Elles incitent à multiplier le nombre de mesures, ce qui n'est pas toujours possible.
- Une solution intermédiaire souvent utilisée dans la littérature scientifique est de donner uniquement les écarts types (avec le nombre de mesures). Ceci permet au lecteur d'approfondir si nécessaire l'étude statistique en fonction de ses besoins.

Ce texte étant une simple introduction aux méthodes d'évaluation des incertitudes, la consultation d'ouvrages spécialisés est par ailleurs indispensable.

BIBLIOGRAPHIE

- Polycopié de PS22 : Mesures physiques
- Polycopié de MT09 : Analyse numérique élémentaire
- Polycopié de SY02 : Méthodes statistiques pour la fabrication et la recherche

Textbook of Quantitative Chemical Analysis par Vogel
Chimie Analytique par Skoog-West-Holler

Cote BUTC: QD 101.2 VOG
Cote BUTC: QD 72-2 SKO

La BUTC propose de nombreux ouvrages de statistiques, par exemple:

Aide Mémoire Pratique de Techniques Statistiques (Ceresta)

Cote BUTC: QA 276 AID