



HAL
open science

La relation de conjugaison et la régression linéaire. Troisième partie : développements

Julien Browaeys, Groupe Irem “ Mesurer En Physique-Chimie ”

► To cite this version:

Julien Browaeys, Groupe Irem “ Mesurer En Physique-Chimie ”. La relation de conjugaison et la régression linéaire. Troisième partie : développements. Le Bulletin de l'Union des Professeurs de Physique et de Chimie, 2021, 115 (1034), pp.551-572. hal-03247161

HAL Id: hal-03247161

<https://hal.science/hal-03247161>

Submitted on 2 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

La relation de conjugaison et la régression linéaire

Troisième partie : développements⁽¹⁾

par **Julien BROWAEYS**

Laboratoire Matière et systèmes complexes (MSC) - CNRS UMR 7057

Université de Paris - 75013 Paris

julien.browaey@u-paris.fr

et **Groupe IREM «Mesurer en physique-chimie»**

Institut de recherche sur l'enseignement des mathématiques (IREM) - 75013 Paris

<https://irem.u-paris.fr>

LA RÉGRESSION linéaire est utilisée comme outil d'analyse en séance de travaux pratiques sur la relation de conjugaison. Dans deux articles précédents, nous avons montré combien cette pratique était critiquable ; puis nous avons proposé une méthode alternative, consistant à calculer la focale pour chaque couple de distances relevées ; cela permet d'en déduire une mesure de la distance focale. Cependant, notre alternative, si elle reste accessible en classe de première, n'est pas optimale. Ici, nous proposons une méthode plus efficace, mais uniquement accessible au niveau universitaire. À l'aide d'une simulation Monte-Carlo, nous montrons, en toute généralité, comment comparer les performances de différentes méthodes d'analyse. Enfin, nous discutons de l'extension possible de notre alternative à l'étude d'autres lois physiques en séance de travaux pratiques.

INTRODUCTION

Deux articles précèdent celui-ci. Dans le premier [1], nous avons montré que la régression linéaire ordinaire était l'outil usuellement choisi pour analyser les résultats expérimentaux en séance de travaux pratiques sur la relation de conjugaison. Cela ne va pas sans poser de nombreux problèmes, tant du point de vue scientifique que didactique. Dans le deuxième [2], nous avons proposé une alternative à la régression linéaire, qui permet de régler la plupart des problèmes soulevés. Elle consiste à calculer pour chaque observation une distance focale. Cet ensemble de distances focales est ensuite représenté, de façon à faire émerger l'idée de variabilité associée à la mesure. La valeur mesurée est égale à la moyenne, et l'évaluation de l'incertitude-type associée est faite à l'aide de l'écart-type divisé par la racine carrée du nombre d'observations. Nous proposons également une procédure pour juger de l'adéquation entre le modèle, c'est-à-dire la relation de conjugaison des lentilles minces, et les observations expérimentales.

(1) **NDLR** : La première partie de cet article est parue dans *Le Bup* n° 1032, mars 2021, p. 347-365, la deuxième partie est parue dans *Le Bup* n° 1033, avril 2021, p. 475-487.

Néanmoins, comme nous l'avons signalé, la méthode alternative proposée n'est pas exempte de défauts. Notre mesure de la distance focale ignore les incertitudes associées aux observations effectuées. Ce choix a été fait dans un souci de simplification pour les élèves de première : il fallait que la méthode reste abordable et explicite. Mais la question demeure de savoir si on ne peut pas mieux faire, en relaxant la contrainte du niveau. C'est pourquoi dans cet article nous nous plaçons au niveau de la licence universitaire, ce qui pourra susciter l'intérêt des étudiant-es préparant le CAPES comme l'agrégation.

Nous nous proposons donc d'exposer ici une méthode plus complète que l'alternative initialement proposée pour la mesure de la distance focale, à l'aide d'une moyenne pondérée. Mais comment prouver que cette nouvelle méthode est plus performante que l'alternative précédente, ou encore que la régression linéaire ordinaire encore omniprésente ? Nous montrerons comment une simulation Monte-Carlo permet de répondre à cette question, en comparant les performances de ces différentes méthodes d'analyse des données. Il est notable que cette comparaison par simulation Monte-Carlo possède un champ d'application qui va bien au-delà de la comparaison des trois méthodes évoquées : elle permet de comparer toute méthode de traitement des données.

Il devient donc possible d'examiner dans quelle mesure on peut substituer nos méthodes d'analyse à la classique régression linéaire dans d'autres travaux pratiques que ceux mettant en œuvre la relation de conjugaison.

1. LA MESURE DE LA DISTANCE FOCALE

On envisage une expérience dans laquelle on projette l'image d'un objet sur un écran au moyen d'une lentille mince. On note O le centre optique de la lentille, A l'intersection du plan objet et de l'axe optique, et A' celle du plan image et de l'axe optique. On utilise l'indice i variant de 0 à $N-1$ pour numérotter les observations réalisées, de façon à ce que la transposition en langage *Python*TM soit plus aisée.

1.1. Critique de l'alternative proposée en classe de première

La méthode alternative à la régression linéaire, précédemment évoquée, consiste à calculer une distance focale pour chaque couple $(\overline{OA_i}, \overline{OA'_i})$ à l'aide de :

$$f'_i = \frac{1}{\frac{1}{\overline{OA'_i}} - \frac{1}{\overline{OA_i}}}.$$

La valeur mesurée de la distance focale est donnée par la moyenne arithmétique $\overline{f'} = \frac{1}{N} \sum_i f'_i$; l'incertitude-type de cette moyenne est donnée par $u(\overline{f'}) = \frac{s_{\text{exp}}}{\sqrt{N}}$, où

s_{exp} est l'écart-type expérimental des f_i' :

$$s_{\text{exp}} = \sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} (f_i' - \overline{f'})^2}.$$

Il s'agit de la procédure classique correspondant à l'évaluation d'une incertitude de type A, au programme de première.

Toutefois il apparaît qu'on n'exploite pas la connaissance expérimentale sur l'incertitude associée aux relevés de \overline{OA} et \overline{OA}' , même si on obtient *in fine* une incertitude-type. Pour résoudre cette situation en apparence paradoxale, il faut revenir au modèle probabiliste sous-jacent à l'évaluation de type A, modèle rarement explicité dans l'enseignement secondaire.

1.1.1. Le modèle sous-jacent à l'évaluation de type A n'est pas valable

Pour décrire la mesure d'une grandeur physique, on la modélise par une variable aléatoire X dont l'espérance μ est par définition la valeur vraie, et dont l'écart-type σ est par définition l'incertitude-type [3]. De la même manière on décrit la répétition de N mesures prises indépendamment, par une suite de variables aléatoires : X_0, X_1, \dots, X_{N-1} . Ces variables aléatoires sont indépendantes et identiquement distribuées (*iid*). En pratique on estime leur écart-type commun σ_X par s_{exp} , l'écart-type expérimental des N valeurs réalisées de X .

Pour approcher aussi bien que possible la valeur vraie, on utilise souvent⁽²⁾ la moyenne arithmétique $M = \frac{1}{N} \sum_{i=0}^{N-1} X_i$ des N mesures répétées. L'espérance de cette moyenne vaut toujours μ , ce qui signifie que faire la moyenne de nos N observations ne crée pas de biais. L'écart-type de cette moyenne σ_M vaut $\frac{\sigma_X}{\sqrt{N}}$. C'est la raison pour laquelle on effectue une moyenne : la variabilité de la valeur mesurée (maintenant la moyenne) décroît avec le nombre d'observations.

Malheureusement, une des hypothèses du modèle sous-jacent à l'évaluation de type A n'est pas vérifiée dans notre expérience. Les N valeurs f_i' , vues comme N variables aléatoires, ne sont pas identiquement distribuées puisque les incertitudes-types associées à chacune de ces valeurs ne sont pas les mêmes. Même si on suppose (pour simplifier) que les incertitudes de \overline{OA}_i et \overline{OA}'_i sont constantes, la loi de propagation des incertitudes implique que celles portant sur les f_i' ne le sont pas [2].

Notre moyenne M a toujours une espérance μ , autrement dit elle réalise bien une

(2) La moyenne arithmétique n'est pas toujours le meilleur estimateur de l'espérance (ici meilleur au sens de non biaisé et qui minimise la variance). Par exemple pour une distribution uniforme, on aura intérêt à prendre la moyenne du minimum et du maximum observés [4]. Quoi qu'il en soit, la pratique courante consiste à utiliser la moyenne.

estimation non biaisée de la distance focale. En effet toutes les observations de f' ont la même espérance, même si elles n'ont pas le même écart-type (incertitude-type). Par contre, l'incertitude-type de M n'est plus donnée par $\sigma_M = \frac{\sigma_X}{\sqrt{N}}$, puisque chacun des f'_i possède une incertitude-type distincte, qu'on ne peut identifier à une valeur unique σ_X . Que vaut alors l'écart-type de la moyenne, et donc l'incertitude-type cherchée ?

1.1.2. L'approche envisagée reste cependant correcte

Pour simplifier, considérons deux variables aléatoires indépendantes X_0 et X_1 , d'espérance commune μ et d'écart-type respectifs σ_{X_0} et σ_{X_1} . Dans ce cas, la variance de la somme est égale à la somme des variances, ce qui entraîne que l'écart-type de $X_0 + X_1$ vaut $\sqrt{\sigma_{X_0}^2 + \sigma_{X_1}^2}$. Il est possible de généraliser dans le cas de la somme de N variables aléatoires, puis à leur moyenne arithmétique qui n'est autre que cette somme divisée par N . En identifiant écart-type et incertitude-type, et en reprenant les notations initiales de notre problème, on obtient :

$$u(\overline{f'}) = \frac{\sqrt{\sum_i u(f'_i)^2}}{N}.$$

Deux cas de figure se présentent :

- ◆ si les incertitudes-types $u(f'_i)$ sont connues, on calcule l'incertitude-type $u(\overline{f'})$ à l'aide de l'expression précédente ;
- ◆ si les incertitudes-types $u(f'_i)$ sont inconnues (tout en sachant qu'elles sont distinctes) on peut continuer d'évaluer $u(\overline{f'})$ par $\frac{s_{\text{exp}}}{\sqrt{N}}$. Ce résultat peut sembler surprenant puisqu'on a levé l'hypothèse de distribution identique des f'_i , mais il est pourtant parfaitement légitime (cf. annexe 1). Ainsi la procédure envisagée au niveau première reste-t-elle valable. Néanmoins comme il s'agit une évaluation statistique, le résultat est d'autant plus variable qu'il y a peu d'observations.

Cette conclusion est rassurante, mais pas entièrement satisfaisante. Car en réalité il est possible de construire un autre estimateur non biaisé, différent de la moyenne arithmétique, et dont l'incertitude-type est inférieure à celle qu'on vient de calculer.

1.2. Proposition au niveau universitaire

Au niveau universitaire il devient possible d'estimer la distance focale par une moyenne pondérée bien choisie, moins soumise à variabilité que la moyenne arithmétique. On réserve cette approche à ce niveau pour deux raisons : d'une part elle exige la détermination des incertitudes des relevés, dont on sait qu'elle est consommatrice d'un temps précieux en séance de travaux pratiques ; d'autre part son calcul, contrairement à la celui de la moyenne, ne se résume plus à une unique instruction, que ce soit avec un tableur ou avec *Python*[™].

Lorsque les incertitudes de chaque observation sont connues (que ce soit par une méthode de type A ou de type B), il est conseillé dans la littérature scientifique de calculer la moyenne pondérée suivante :

$$\langle f' \rangle = \frac{\sum_{i=0}^{N-1} w_i f'_i}{\sum_{i=0}^{N-1} w_i},$$

où les poids w_i sont égaux à l'inverse du carré de l'incertitude-type de f'_i : $w_i = \frac{1}{u(f'_i)^2}$ (cf. annexe 2 pour une justification détaillée). Cette pondération est celle qui minimise l'incertitude-type $u(\langle f' \rangle)$, qui vaut alors :

$$u(\langle f' \rangle) = \frac{1}{\sqrt{\sum_i \frac{1}{u(f'_i)^2}}}.$$

D'autres pondérations existent, mais elles font toujours l'objet de discussions [5]. C'est pourquoi il nous semble suffisant de se limiter à la version indiquée au niveau licence.

1.3. Comparaison des deux méthodes alternatives

Comme nous l'avons vu dans la procédure proposée au niveau lycée, il n'est pas nécessaire de connaître différentes valeurs de $u(f'_i)$ pour estimer la distance focale et son incertitude-type associée. À l'inverse, notre procédure au niveau universitaire exige de connaître les valeurs de $u(f'_i)$. Pour évaluer l'amélioration qu'elle confère, nous nous proposons de reprendre de notre précédent article [2] les observations expérimentales relevées : \overline{OA}_i , \overline{OA}'_i , $u(\overline{OA}_i)$, $u(\overline{OA}'_i)$.

Pour accéder aux incertitudes $u(f'_i)$, on peut soit employer une simulation Monte-Carlo comme dans notre précédent article [2], ou procéder directement par la formule de propagation de l'incertitude qui est désormais accessible au niveau universitaire :

$$u(f'_i) \simeq \sqrt{\left(\frac{\partial f'_i}{\partial \overline{OA}_i} u(\overline{OA}_i)\right)^2 + \left(\frac{\partial f'_i}{\partial \overline{OA}'_i} u(\overline{OA}'_i)\right)^2} = f_i'^2 \sqrt{\frac{(u(\overline{OA}_i))^2}{\overline{OA}_i^4} + \frac{(u(\overline{OA}'_i))^2}{\overline{OA}'_i^4}}.$$

En pratique, la différence entre la moyenne pondérée et la moyenne arithmétique reste modérée et comparable aux incertitudes, comme attendu. C'est ce que montrent les résultats des calculs⁽³⁾ du tableau 1 (cf. page ci-après), sur la base des observations expérimentales déjà réalisées [2]. L'amélioration en termes d'incertitude reste limitée parce que les incertitudes-types $u(f'_i)$ sont relativement proches les unes des autres. Cela justifie *a posteriori* l'approche proposée au lycée.

(3) Les données, les calculs effectués et les figures réalisées sont disponibles en ligne, dans une feuille de calcul *Google Spreadsheet* (<https://bit.ly/2E68LWT>) qu'il est possible d'importer au format *Microsoft Excel* ou *LibreOffice* ; et dans un notebook *Google Colab* (<https://bit.ly/2Fn2xdQ>) qu'il est possible de récupérer comme notebook *Jupyter* ou fichier *Python*[™].

Méthode	Valeur mesurée	Incertitude-type associée
Moyenne arithmétique	$\overline{f'} = 20,210 \text{ cm}$	$u(\overline{f'}) = 0,023 \text{ cm}$
Moyenne pondérée	$\langle \overline{f'} \rangle = 20,228 \text{ cm}$	$u(\langle \overline{f'} \rangle) = 0,018 \text{ cm}$

Tableau 1 - Comparaison des résultats de mesure à l'aide de la moyenne pondérée et de la moyenne arithmétique. Les expressions utilisées sont celles des sous-paragraphes 1.1.2. et 1.2.

Dans notre précédent article [2], nous avons obtenu une valeur de l'incertitude-type de 0,015 cm. Il ne faut pas s'étonner de la différence avec la valeur $u(\overline{f'})$ ici calculée à l'aide des $u(f'_i)$: l'évaluation était précédemment statistique, donnée par s_{exp}/\sqrt{N} ; elle est donc soumise à des fluctuations d'autant plus importantes que le nombre d'observations est faible. La comparaison entre l'estimation statistique de l'incertitude (à partir des f'_i) et l'estimation à partir des $u(f'_i)$ est essentielle pour la validation du modèle, comme nous allons le voir.

2. LA VALIDATION DU MODÈLE

Outre la question de la mesure de la distance focale, il se pose celle, distincte, de la validité de la relation de conjugaison des lentilles minces. Il faut bien comprendre que cette relation de conjugaison est une modélisation d'une situation expérimentale, dans laquelle certains phénomènes sont négligés, comme par exemple l'épaisseur de la lentille, ou la diffraction [2]. Connaissant nos incertitudes, est-il possible, ou non, d'invalider le modèle choisi ? Autrement dit, nos mesures expérimentales et le modèle sont-ils compatibles ? Cette question est d'autant plus cruciale que la mesure de la distance focale proposée à la section précédente n'a de sens qu'à la condition que le modèle de la lentille mince convienne.

Pour y répondre, on avait envisagé [2] de comparer *successivement* chacune des observations ($f'_i, u(f'_i)$) et la valeur finale, à savoir $\overline{f'}$. Ce faisant, on suit un raisonnement en partie circulaire puisque la valeur de $\overline{f'}$ est issue de la moyenne des différents f'_i : pour s'en rendre compte il suffit de considérer le cas $N=2$. Un autre problème réside dans le nombre N des comparaisons. Même s'il y a compatibilité globale, plus on augmente N , plus la probabilité de voir une observation fluctuer au-delà du seuil fixé à l'avance augmente. Il faudrait donc modifier le seuil en fonction de N , par exemple réaliser une correction de Bonferroni [6]. C'est inabordable au lycée.

En réalité, il faut voir si les N valeurs mesurées de f'_i , accompagnées de leur incertitude-type associée $u(f'_i)$, sont *simultanément* compatibles entre elles. Vu de manière moderne, dans le contexte du GUM [7-8], il est recommandé de calculer la

grandeur suivante⁽⁴⁾, carré du ratio de Birge (cf. annexe 3) :

$$R_{\text{exp}}^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} \left(\frac{f_i' - \langle f' \rangle}{u(f_i')} \right)^2.$$

Cette expression compare les fluctuations observées $f_i' - \langle f' \rangle$ à celles estimées par $u(f_i')$. Si les mesures sont compatibles, on s'attend à ce que chacun des N termes de la somme vaille à peu près 1. Parce que $\langle f' \rangle$ est estimé à partir des observations $(f_i', u(f_i'))$, la somme devrait valoir en réalité environ $N - 1$, ou, ce qui revient au même, R_{exp}^2 devrait être de l'ordre de l'unité [9]. Si sa valeur est notablement plus importante, on doit conclure à l'incompatibilité du modèle et des observations (cf. annexe 3 et [2] pour une discussion à ce sujet).

Il reste à déterminer un seuil pour avoir un critère de décision. Moyennant certaines approximations usuellement justifiées, on considère que $(N - 1) R_{\text{exp}}^2$ est la réalisation d'une variable aléatoire $(N - 1) R^2$ qui suit une loi de probabilité du χ^2 à $N - 1$ degrés de libertés [9] :

$$(N - 1) R^2 \sim \chi_{N-1}^2.$$

Un test statistique⁽⁵⁾ est alors envisageable, en se fixant un risque d'erreur p , typiquement 0,05. Le seuil R_{seuil}^2 est alors fixé par l'équation suivante :

$$\mathbb{P}(R^2 \geq R_{\text{seuil}}^2) = p.$$

Si les mesures sont bien compatibles, il arrive avec une probabilité p faible que des fluctuations intenses se produisent. Dans ce cas, on sera amené à rejeter à tort la compatibilité. Mais comme une incompatibilité est susceptible de produire un ratio R^2 élevé, on espère que cette procédure permette de la détecter.

Pour mieux comprendre, on peut considérer l'exemple d'un jeu de pile ou face, dont on suppose *a priori* qu'il n'est pas truqué. Une personne qui joue et qui perd dix fois de suite sait que la probabilité que cela arrive est faible ($1/2^{10} = 1/1024$), mais pas nulle. C'est donc possiblement une fluctuation statistique. Mais cette personne va néanmoins arrêter de jouer, car elle soupçonne que le jeu est truqué.

La figure 1 (cf. page ci-après) représente R_{seuil}^2 en fonction du nombre de mesures N , pour deux niveaux de risque p donnés. La plupart des logiciels classiques sont en mesure d'exprimer R_{seuil}^2 en fonction du seuil choisi. Le tableau 2 (cf. page ci-après) suivant indique comment procéder avec un tableur ou en langage *Python*TM.

- (4) Il ne faut pas la confondre avec le coefficient de détermination R^2 d'une régression linéaire.
 (5) Pour étayer ce choix, des connaissances élémentaires en statistique inférentielle élémentaire sont nécessaires. Dans notre expérience, les cursus de licence de physique incluent rarement des cours de statistique. Ce ne sont pas les difficultés mathématiques qui sont en cause puisqu'à l'université, que ce soit en biologie, en médecine, en psychologie ou en économie, de tels cours sont presque toujours au programme de licence.

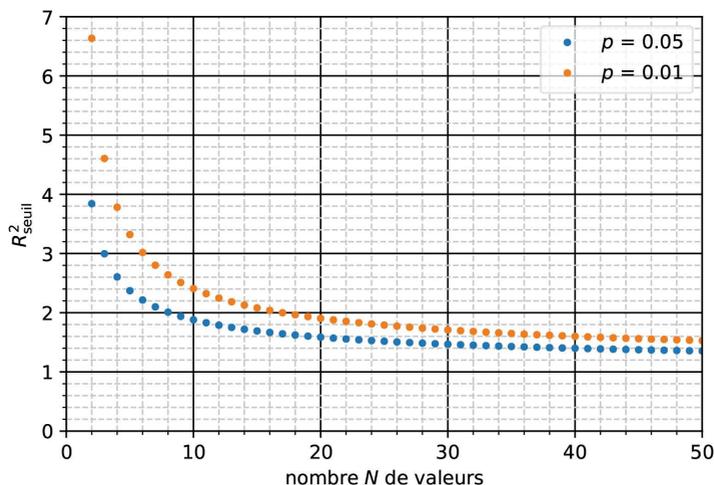


Figure 1 - Seuil de R^2_{exp} au-delà duquel on conclut à l'incompatibilité du modèle et des données. Par exemple, pour $N=3$ valeurs et $p=0,05$, le seuil vaut environ 3.

Logiciel	Expression pour obtenir le seuil R^2_{seuil} au risque p
Microsoft Excel (français) LibreOffice (français)	= KHIDEUX.INVERSE(p ; $N-1$)/($N-1$) ou ⁽⁵⁾ = LOI.KHIDEUX.INVERSE.DROITE(p ; $N-1$)/($N-1$)
Microsoft Excel (anglais) LibreOffice (anglais) Google Spreadsheets	= CHIINV(p ; $N-1$)/($N-1$) ou ⁽⁵⁾ = CHISQ.INV.RT(p ; $N-1$)/($N-1$)
Python™	<pre>from scipy.stats import chi2 # importe la fonction chi2 R2_seuil = chi2.isf(p, N-1)/(N-1) # calcule le seuil</pre>

Tableau 2 - Instructions à utiliser pour calculer le seuil R^2_{seuil} en fonction du nombre de mesures N et du niveau de risque p choisi.

Pour résumer, lorsque R^2_{exp} dépasse le seuil R^2_{seuil} qu'on a calculé à partir de N et du niveau de risque choisi (souvent 0,05), on rejette l'hypothèse de compatibilité entre les données et le modèle. Avec les données rapportées dans notre précédent article [2], on trouve (cf. note 3) : $R^2_{\text{exp}} = 0,820$. Puisque $N = 7$, la valeur limite vaut $R^2_{\text{seuil}} = 2,10$ au risque $p = 0,05$. Ainsi conclut-on que nos mesures peuvent être considérées comme compatibles avec la relation de conjugaison des lentilles minces.

(6) L'usage dans *Microsoft Excel* de la fonction `KHIDEUX.INVERSE` (ou `CHIINV`) n'est plus recommandé depuis 2007, bien qu'elle soit toujours opérationnelle ; il n'est pas sûr que cela continue dans les futures versions du logiciel.

3. LA COMPARAISON DES DIFFÉRENTES MÉTHODOLOGIES

Les critères jusqu'ici développés pour comparer les performances des différentes méthodes d'analyse des données étaient de trois types : (1) théoriques, en convoquant les lois des probabilités, (2) dogmatiques, en invoquant les pratiques scientifiques, ou (3) didactiques, en se plaçant à portée des élèves ou des étudiant-es. On peut y ajouter un critère supplémentaire, de type calculatoire, s'appuyant sur une simulation Monte-Carlo déjà abordée dans notre précédent article [2]. Il est en effet possible d'évaluer par une expérience numérique l'efficacité relative de différentes méthodes d'analyse des données. Il nous paraît souhaitable de s'en saisir dans un contexte d'enseignement, car il s'agit d'une procédure relativement simple permettant une analyse critique, ce qui constitue une part importante de l'activité expérimentale.

3.1. Le principe de la simulation Monte-Carlo

Pour introduire cette procédure, nous allons traiter à titre d'exemple les trois méthodes de mesure de la distance focale évoquées dans cet article et dans les précédents [1-2] : la régression linéaire ordinaire, la moyenne arithmétique et la moyenne pondérée. Il n'en reste pas moins que cette procédure peut s'appliquer à un ensemble de problèmes bien plus large, débordant du champ disciplinaire de la physique.

On suppose qu'on dispose d'une lentille de focale donnée (disons 20 cm), que la relation de conjugaison des lentilles minces est vraie, et que l'on impose sept distances \overline{OA}_i , typiques (proches de celles effectivement réalisées) : on peut en déduire sept distances \overline{OA}'_i . Nous obtenons donc sept couples de valeurs $(\overline{OA}_i, \overline{OA}'_i)$ exactes. De chacun, on peut déduire une distance focale valant *exactement* 20 cm.

Pour modéliser le processus de relevé des valeurs de $(\overline{OA}_i, \overline{OA}'_i)$, on va rajouter à ces valeurs une variable aléatoire simulée numériquement, dont la dispersion correspond à la variabilité estimée de nos observations. Par exemple, si on a estimé que l'observation de \overline{OA}_i se faisait à $\pm 0,1$ cm près, on va rajouter à chacune des valeurs de \overline{OA}_i un nombre aléatoire tiré d'une distribution (par exemple) uniforme centrée sur 0 cm, et de demi-étendue 0,1 cm. On procède de même pour les \overline{OA}'_i (voir [2] pour une discussion sur l'évaluation de l'incertitude de \overline{OA}'_i). À partir de ces sept couples de valeurs $(\overline{OA}_i, \overline{OA}'_i)$ qui simulent une expérience, on calcule la focale par les trois différentes méthodes évoquées : on obtient trois valeurs de la focale, *a priori* toutes légèrement différentes de 20 cm. On recommence un grand nombre de fois cette procédure, de façon à accumuler un grand nombre d'expériences simulées.

Pour chacune des trois méthodes de mesure envisagées, on obtient alors une collection de distances focales, dont on peut tracer l'histogramme. Il est possible d'extraire les indicateurs de position et de dispersion que sont la moyenne et l'écart-type. Pour que la méthode soit juste, il faut que la moyenne soit aussi proche que possible

des 20 cm initiaux ; si l'écart est notable, on constate un biais qu'il faudra corriger, le cas échéant, en modifiant la méthode. Pour que la méthode soit fidèle, il faut que l'écart-type soit aussi faible que possible.

3.2. Les résultats de la simulation Monte-Carlo

Nous avons réalisé (cf. note 3) cinquante mille simulations numériques d'expérience en langage *Python*TM. La figure 2 représente en superposition les histogrammes de densité de fréquence des résultats issus des trois méthodes ; le tableau 3 résume les résultats à l'aide de la moyenne et de l'écart-type.

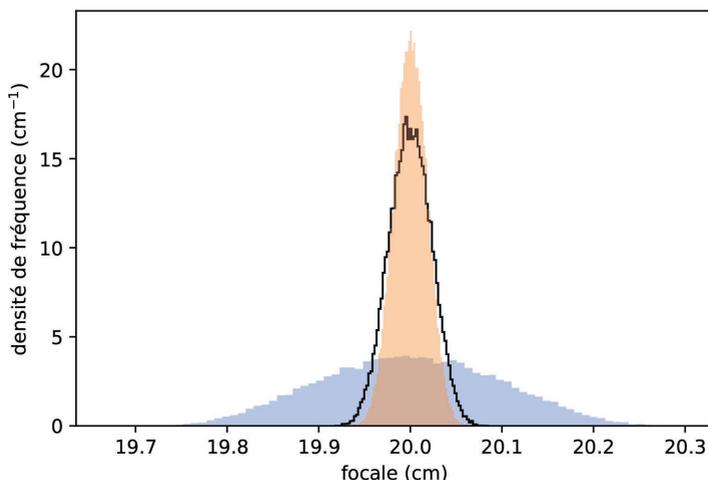


Figure 2 - Résultat de la simulation Monte-Carlo pour la mesure de la focale (cinquante mille essais) - bleu : régression linéaire ordinaire ; noir : moyenne arithmétique ; orange : moyenne pondérée.

Méthode	Valeur moyenne (cm)	Écart-type (cm)
Régression linéaire ordinaire	19,9998	0,0959
Moyenne arithmétique	19,9992	0,0231
Méthode pondérée	19,9999	0,0180

Tableau 3 - Résultats de la simulation Monte-Carlo visant à comparer⁽⁷⁾ les trois méthodes d'analyse.

On constate d'abord que les trois méthodes sont justes, c'est-à-dire non biaisées. En effet, les valeurs moyennes des focales simulées demeurent extrêmement proches

(7) On écrit ici quatre chiffres derrière la virgule de façon à mettre en évidence les différentes valeurs.

de la valeur initiale choisie de 20 cm. Autrement dit la différence entre les valeurs moyennes et 20 cm est négligeable devant l'écart-type, c'est-à-dire l'incertitude-type.

Ensuite, on relève que la méthode de la moyenne pondérée est celle qui présente la dispersion la plus faible, ce qui légitime son utilisation préférentielle. Néanmoins, la méthode par la moyenne arithmétique simple est quasiment aussi efficace. Nous l'avons déjà remarqué lors de la comparaison des moyennes arithmétique et pondérée ; mais cette fois-ci, cette constatation est issue d'une simulation numérique d'expérience et non plus de calculs de probabilité. La régression linéaire ordinaire, quant à elle, est de loin la moins performante des trois méthodes, ce qui achève de la discréditer dans ce cadre.

Enfin, on observe que les évaluations d'incertitude-type prédites par chaque méthode sont bien cohérentes⁽⁸⁾ avec celles obtenues par la simulation (cf. tableau 4).

Méthode	Incrtitude-type prédite (cm)	Incrtitude-type observée (cm)
Moyenne arithmétique	0,0225	0,0231
Moyenne pondérée	0,0181	0,0180

Tableau 4 - Comparaison (cf. note 8) entre l'incrtitude-type prédite et celle observée dans la simulation Monte-Carlo.

Toute méthode de traitement des données peut ainsi être jaugée à l'aune d'une simulation Monte-Carlo. Que l'incrtitude résulte à la fois de l'opérateur, des instruments, des observations relève presque d'un lieu commun. Nous avons ici mis en évidence qu'elle dépend également des choix opérés dans le traitement des données.

4. LA VIABILITÉ DES MÉTHODES ALTERNATIVES DANS D'AUTRES CONTEXTES

Les méthodes proposées dans cet article et le précédent, à savoir la moyenne arithmétique et la moyenne pondérée forment une alternative à la régression linéaire ordinaire, dans le contexte de travaux pratiques sur la relation de conjugaison des lentilles minces. Il semble possible d'exploiter ces méthodes dans d'autres situations où l'on souhaite mesurer un paramètre physique issu d'une loi reliant deux grandeurs. À titre d'exemple, le tableau 5 (cf. page ci-après) énonce comment on pourrait procéder dans plusieurs cas.

(8) Pour s'en assurer il faudrait évaluer l'écart-type des incrtitudes-type observées dans plusieurs simulations Monte-Carlo et également l'écart-type de l'incrtitude-type prédite.

Loi physique	Relation utilisée	Paramètre recherché	Variables observées	Valeurs à exploiter
Loi de conjugaison	$\frac{1}{\overline{OA'}} - \frac{1}{\overline{OA}} = \frac{1}{f'}$	f'	$\overline{OA}, \overline{OA'}$	$\frac{\overline{OA'} \cdot \overline{OA}}{\overline{OA} - \overline{OA'}}$
Loi d'Ohm	$U = RI$	R	U, I	U/I
Loi de Snell-Descartes	$\sin i = n \sin r$	n	i, r	$\frac{\sin i}{\sin r}$
Loi du pendule	$T = 2\pi \sqrt{\frac{L}{g}}$	g	T, L	$\frac{4\pi^2 L}{T^2}$
Loi de l'hydrostatique	$\Delta P = \rho g \Delta z$	ρg ou ρ ou g	$\Delta P, \Delta z$	$\Delta P / \Delta z$

Tableau 5 - Différentes situations où les mesures sont analysables par la méthode des moyennes.

4.1. Une limitation, la division par zéro

On remarque que les valeurs à exploiter se présentent toujours sous forme d'un rapport de la forme $\mathcal{E} = \mathcal{N}/\mathcal{D}$ (où \mathcal{E} représente la valeur à exploiter, \mathcal{N} son numérateur et \mathcal{D} son dénominateur). Si le domaine d'observation des variables est tel que \mathcal{N} et \mathcal{D} s'approchent de zéro, on risque d'avoir un problème. En effet :

$$u(\mathcal{E}) \simeq \mathcal{E} \sqrt{\frac{u(\mathcal{N})^2}{\mathcal{N}^2} + \frac{u(\mathcal{D})^2}{\mathcal{D}^2}},$$

ce qui signifie que l'incertitude sur la valeur de \mathcal{E} pourrait devenir très grande.

Dans le cas de la relation de conjugaison, le dénominateur (et donc le numérateur) n'est jamais susceptible de s'annuler puisque $|\overline{OA} - \overline{OA'}| \geq 4f'$ pour un objet et une image réelle. Mais dans les autres cas, il faut prendre ses précautions. En pratique, cela signifie que pour déterminer un paramètre d'une loi reliant deux grandeurs, qu'on suppose vraie, on a intérêt, autant que possible, à utiliser les valeurs de ces grandeurs qui sont éloignées de zéro.

Ici encore, il convient de réaliser une étude par simulation Monte-Carlo pour voir quelle est la méthode la plus performante : la régression linéaire peut éventuellement se révéler un meilleur choix dans le contexte donné.

4.2. D'autres limitations

Il arrive parfois que le modèle à utiliser pour décrire les observations ne soit pas linéaire, mais affine, bien que cela ne se voie pas dans l'expression des lois physiques sous-jacentes. Deux exemples permettent de le mettre en évidence :

- ◆ $U = RI$ est une loi physique linéaire, avec un unique paramètre R reliant les deux variables U et I . Cependant, si on utilise des multimètres dont le zéro est mal calibré, il est nécessaire d'effectuer une modélisation par une loi affine.
- ◆ $\Delta P = \rho g \Delta z$ est également une loi physique linéaire ; mais il y a toujours une difficulté à calculer exactement une différence, à cause de la résolution des instruments par exemple ; aussi paraît-il préférable de modéliser la loi de l'hydrostatique par : $P = P_0 + \rho g z$, une loi affine donc.

Dans les cas évoqués, une régression linéaire ordinaire devient préférable puisqu'il faut estimer deux paramètres indépendants, ayant chacun un sens dans le cadre d'un modèle physique.

CONCLUSION

Cet article vient clore une réflexion en trois parties sur une situation expérimentale d'enseignement relative à la loi de conjugaison des lentilles minces. Dans un premier article, nous avons montré, à travers une analyse des manuels, que la pratique sco-

laire utilisait la régression linéaire comme outil d'analyse, tant pour évaluer la distance focale que pour justifier la validité de la relation de conjugaison. Nous y avons détaillé les multiples raisons pour lesquelles cet usage nous paraît inadapté. Dans un deuxième article, nous avons proposé une alternative qui s'inscrit dans le cadre du programme de première et qui évite les écueils de la régression linéaire. Elle est réalisable aussi bien à l'aide d'un tableur qu'avec un code Python™. Dans ce troisième article, pour résoudre certaines limitations de l'alternative, nous étendons notre proposition en nous plaçant au niveau de la licence universitaire. Nous montrons comment une simulation Monte-Carlo peut servir à comparer différentes méthodes d'analyse des données, conformément à la pratique scientifique contemporaine. Le langage Python™ permet de réaliser aisément de telles simulations. Enfin nous explorons en conclusion l'applicabilité de l'alternative proposée à d'autres situations expérimentales que l'optique géométrique. Dans les trois articles, nous avons joint des fichiers d'exemples.

Il ne faudrait pas conclure de nos articles que la régression linéaire est systématiquement à bannir de toute séance de travaux pratiques. Nous avons simplement essayé de montrer que dans un cas particulier, une alternative était susceptible de la remplacer avantageusement. L'extension du champ d'application de notre proposition reste évidemment à étudier. Pour cela, comme nous nous sommes efforcés de le faire, il faudra mobiliser des arguments liés à l'élève (ce qu'il est possible d'enseigner), à l'institution (les programmes), à la communauté scientifique (ses usages), à la théorie mathématique (la statistique). Parce que de nombreux textes sur l'enseignement de la physique sont plus prescriptifs que discursifs, les raisons des choix didactiques qui ont été faits se sont perdues avec le temps. Les pratiques scolaires évoluent alors parfois jusqu'à devenir des rituels. Notre ambition est de contribuer à ce que la communauté les interroge pour élaborer de nouvelles manières de faire.

REMERCIEMENTS

Les travaux collectifs du groupe IREM «Mesurer en physique-chimie» sont à l'origine de cet article. L'auteur remercie plus particulièrement Nicolas Décamp, Tristan Beau, Peter Birkenwald et Alexandra Chauvin pour leur relecture critique et leurs judicieuses suggestions.

BIBLIOGRAPHIE ET NETOGRAPHIE

- [1] J. Browaeys et Groupe IREM «Mesurer en physique-chimie», «La relation de conjugaison et la régression linéaire - Première partie : critique», *Bull. Un. Prof. Phys. Chim.*, vol. 115, n° 1032, p. 347-365, mars 2021.

- [2] J. Browaeys, T. Beau et Groupe IREM « Mesurer en physique-chimie », « La relation de conjugaison et la régression linéaire – Deuxième partie : alternative », *Bull. Un. Prof. Phys. Chim.*, vol. 115, n° 1033, p. 475-487, avril 2021.
- [3] F. Grégis, « La valeur de l'incertitude : l'évaluation de la précision des mesures physiques et les limites de la connaissance expérimentale », Thèse de doctorat, Université Paris-Diderot (Paris 7), 2016.
Accessible à : <https://bit.ly/3oaQ8UX>
- [4] Z. L. Warsza and S. Kubisa, "Mid-range as estimator of samples from uniform and flatten-Gaussian distribution", *Proceedings of 20th IMEKO TC4 Symposium*, Benevento, Italy, 15-17 September 2014.
Accessible à : <http://bit.ly/3uLJ31k>
- [5] H. Huang, "A new method for estimating consensus values in interlaboratory comparisons", *Metrologia*, n° 55, p. 106-113, 2018.
<https://doi.org/10.1088/1681-7575/aaa170>
- [6] L. Wasserman, *All of Statistics : A Concise Course in Statistical Inference*, New York, Springer-Verlag, 2004.
<https://doi.org/10.1007/978-0-387-21736-9>
- [7] JCGM, *Évaluation des données de mesure – Guide pour l'expression de l'incertitude de mesure (GUM)*, 2008.
Accessible à : <https://bit.ly/311prNit>
- [8] R. N. Kacker, R. Kessel and K.-D. Sommer, "Assessing Differences Between Results Determined According to the Guide to the Expression of Uncertainty in Measurement", *Journal of Research of the National Institute of Standards and Technology*, n° 115, p. 453-459, 2010.
<https://doi.org/10.6028/jres.115.031>
- [9] R. N. Kacker, A. Forbes, R. Kessel and K.-D. Sommer, "Classical and Bayesian interpretation of the Birge test of consistency and its generalized version for correlated results from interlaboratory evaluations", *Metrologia*, n° 45, p. 257-264, 2008.
<https://doi.org/10.1088/0026-1394/45/3/001>
- [10] A. H. Joarder and M. H. Omar, "The Mathematical Expectation of Sample Variance: A General Approach", *Bull. Malays. Math. Sci. Soc.*, n° 38, p. 635-646, 2015, DOI : 10.1007/s40840-014-0040-x
Accessible à : <https://bit.ly/31jTUKn>
- [11] R. T. Birge, "The calculation of errors by the method of least squares", *Physical Review*, n° 40, p. 207-227, 1932.
<https://doi.org/10.1103/PhysRev.40.207>

- [12] Wikipédia, article “Inverse-variance weighting”.
Accessible à : <https://bit.ly/3fppcdP>, page consultée le 7 mai 2021.
- [13] Wikipédia, article « Moyenne généralisée ».
Accessible à : <https://bit.ly/2EwBrZH>, page consultée le 7 mai 2021.
- [14] B. N. Taylor, W. H. Parker and D. N. Langenberg, “Determination of e/h , Using Macroscopic Quantum Phase Coherence in Superconductors : Implications for Quantum Electrodynamics and the Fundamental Physical Constants”, *Reviews of Modern Physics*, n° 41, p. 375–496, 1969.
<https://doi.org/10.1103/RevModPhys.41.375>
- [15] M. Tanabashi and *al.* (Particle Data Group), “Review of Particle Physics”, *Physical Review D*, n° 98, p. 16, 2018. Une version différemment formatée est accessible à :
<https://bit.ly/3iLFEaJ>
- [16] M. G. Cox, “The evaluation of key comparison data: An introduction”, *Metrologia*, n° 39, p. 587–588, 2002.
<https://doi.org/10.1088/0026-1394/39/6/9>
- [17] CCQM, *CCQM Guidance note: estimation of a consensus KCRV and associated degrees of equivalence*, version 10, 2013.
Accessible à <https://bit.ly/2RNtUMv>

Annexe 1

Moyenne arithmétique

La moyenne arithmétique est une variable aléatoire calculée à partir de N variables aléatoires X_0, X_1, \dots, X_{N-1} représentant les N mesures prises indépendamment :

$$M = \frac{\sum_{i=0}^{N-1} X_i}{N}.$$

Il s'agit d'un estimateur de l'espérance μ supposée commune des N variables aléatoires X_i . En effet :

$$\mathbb{E}(M) = \frac{\sum_{i=0}^{N-1} \mathbb{E}(X_i)}{N} = \frac{\sum_{i=0}^{N-1} \mu}{N} = \mu.$$

Lorsqu'on calcule une moyenne expérimentale, on obtient une réalisation de la variable aléatoire M . De ce qui précède on comprend qu'elle approche la valeur μ . Mais évidemment, le résultat est soumis à des fluctuations statistiques, dont on estime l'intensité à l'aide de la variance.

La variance de l'estimateur M vaut :

$$\mathbb{V}(M) = \frac{\mathbb{V}\left(\sum_{i=0}^{N-1} X_i\right)}{N^2} = \frac{\sum_{i=0}^{N-1} \mathbb{V}(X_i)}{N^2}$$

(on a utilisé le fait que la variance est une forme quadratique et que les X_i étaient indépendants). Sachant que l'écart-type de la moyenne est $\sigma_M = \sqrt{\mathbb{V}(M)}$, on retrouve l'expression proposée au sous-paragraphe 1.1.2. dans le contexte de la mesure d'une distance focale :

$$\sigma_M = \frac{\sqrt{\sum_{i=0}^{N-1} \mathbb{V}(X_i)}}{N}.$$

Mais comment évaluer $\mathbb{V}(M)$ si on ne connaît pas les $\sigma_{X_i} = \sqrt{\mathbb{V}(X_i)}$? On utilise un estimateur de la variance, comme on a fait pour estimer l'espérance μ . L'estimateur classique de la variance est :

$$S^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (X_i - M)^2.$$

Il est possible de démontrer [10] que l'espérance de cet estimateur de la variance est égal à la somme des variances des X_i divisée par N :

$$\mathbb{E}(S^2) = \frac{\sum_i \mathbb{V}(X_i)}{N}.$$

On en déduit donc que :

$$\mathbb{V}(M) = \frac{\mathbb{E}(S^2)}{N}.$$

Lorsqu'on calcule une variance expérimentale (s_{exp}^2) on obtient une réalisation

de la variable aléatoire S^2 . De ce qui précède on comprend qu'elle approche $\mathbb{E}(S^2)$. Mais évidemment le résultat est soumis à des fluctuations statistiques. Cela permet néanmoins d'estimer la variance de la moyenne, et donc l'écart-type de la moyenne :

$$\mathbb{V}(M) \simeq \frac{s_{\text{exp}}^2}{N} \Rightarrow \sigma_M \simeq \frac{s_{\text{exp}}}{\sqrt{N}}.$$

Annexe 2

Moyenne pondérée

L'idée d'utiliser une moyenne pondérée pour agréger différentes mesures (avec leurs incertitudes respectives) est parfois attribuée à un article séminal de Birge publié en 1932 [3, 11]. Celui-ci avait pour objectif de combiner différents résultats pour déterminer au mieux des constantes fondamentales en physique. Cette analyse est contemporaine de l'invention des concepts de la statistique (Fisher publie *The Design of Experiments* en 1935) comme ceux de la mécanique quantique (Dirac publie *The Principles of Quantum Mechanics* en 1930). Elle est donc relativement récente dans l'histoire des sciences. On comprend dès lors la difficulté de sa transposition didactique.

En reprenant les notations de l'annexe 1, la moyenne pondérée d'une série de variables aléatoires indépendantes X_i d'espérance commune μ est définie par :

$$\langle X \rangle = \frac{\sum_{i=0}^{N-1} w_i X_i}{\sum_{i=0}^{N-1} w_i},$$

où les $w_i > 0$ sont les poids. La moyenne arithmétique est donc un cas particulier de la moyenne pondérée, avec $w_i = 1$.

En utilisant la linéarité de l'espérance, on montre que :

$$\mathbb{E}(\langle X \rangle) = \frac{\sum_{i=0}^{N-1} w_i \mathbb{E}(X_i)}{\sum_{i=0}^{N-1} w_i} = \frac{\sum_{i=0}^{N-1} w_i \mu}{\sum_{i=0}^{N-1} w_i} = \mu.$$

Comme pour la moyenne arithmétique, lorsqu'on calcule une moyenne expérimentale pondérée on s'attend à ce qu'elle approche la valeur μ . Le résultat est là encore soumis à des fluctuations statistiques dont on estime l'intensité à l'aide de la variance.

En utilisant le fait que la variance est une forme quadratique et l'indépendance des X_i (et donc des $w_i X_i$, les w_i étant fixés) on peut exprimer $\mathbb{V}(\langle X \rangle)$:

$$\mathbb{V}(\langle X \rangle) = \frac{\mathbb{V}\left(\sum_{i=0}^{N-1} w_i X_i\right)}{\left(\sum_{i=0}^{N-1} w_i\right)^2} = \frac{\sum_{i=0}^{N-1} w_i^2 \mathbb{V}(X_i)}{\left(\sum_{i=0}^{N-1} w_i\right)^2}.$$

Étant données les $\mathbb{V}(X_i)$, de toutes les pondérations existantes, il en existe une qui minimise $\mathbb{V}(\langle X \rangle)$ [12] :

$$w_i = \frac{1}{\mathbb{V}(X_i)}.$$

En choisissant cette pondération, une observation très incertaine, c'est-à-dire avec une grande variance, va être affectée d'un poids très faible. À l'inverse, une observation pour laquelle l'incertitude est faible aura un poids plus élevé.

Ce choix optimal de pondération effectué, on peut exprimer à nouveau la variance de la moyenne pondérée :

$$\mathbb{V}(\langle X \rangle) = \frac{\sum_{i=0}^{N-1} \frac{1}{\mathbb{V}(X_i)}}{\left(\sum_{i=0}^{N-1} \frac{1}{\mathbb{V}(X_i)} \right)^2} = \frac{1}{\sum_{i=0}^{N-1} \frac{1}{\mathbb{V}(X_i)}}.$$

Chaque nouvelle mesure augmente le dénominateur de l'expression précédente. C'est pourquoi toute nouvelle mesure diminue l'incertitude de la moyenne pondérée. Cette propriété souhaitable n'est pas nécessairement observée avec la moyenne arithmétique classique. Dans le même ordre d'idée, on observe également que $\mathbb{V}(\langle X \rangle)$ reste inférieure à la plus petite des variances de X_i :

$$\frac{1}{\mathbb{V}(\langle X \rangle)} = \sum_i \frac{1}{\mathbb{V}(X_i)} > \frac{1}{\min_i \mathbb{V}(X_i)} \Rightarrow \mathbb{V}(\langle X \rangle) < \min_i \mathbb{V}(X_i).$$

Traduit en termes de mesure, cela signifie que l'incertitude sur $\langle X \rangle$ reste inférieure à la plus petite des incertitudes des X_i .

Puisque l'écart-type de la moyenne pondérée vaut $\sigma_{\langle X \rangle} = \sqrt{\mathbb{V}(\langle X \rangle)}$, on retrouve l'expression proposée au sous-paragraphe 1.2. dans le contexte de la mesure d'une distance focale :

$$\sigma_{\langle X \rangle} = \frac{1}{\sqrt{\sum_{i=0}^{N-1} \frac{1}{\mathbb{V}(X_i)}}}.$$

L'utilisation de la moyenne arithmétique non pondérée $M = \frac{\sum_{i=0}^{N-1} X_i}{N}$ est sous-optimale parce que son écart-type σ_M est plus grand que $\sigma_{\langle X \rangle}$. Il est possible d'interpréter cette inégalité en remarquant qu'une moyenne harmonique de nombres positifs est toujours inférieure à une moyenne arithmétique, en exploitant la convexité de la fonction $x \mapsto 1/x$ [13] ; ici $N\mathbb{V}(\langle X \rangle)$ est la moyenne harmonique des $\mathbb{V}(X_i)$, tandis que $N\mathbb{V}(M)$ est la moyenne arithmétique des $\mathbb{V}(X_i)$. Notons qu'en toute rigueur cette inégalité n'est pas stricte : les deux moyennes sont égales si et seulement si les variances $\mathbb{V}(X_i)$ (et donc les incertitudes-types) sont toutes égales.

Toutes ces considérations sont parfaitement compréhensibles par des étudiants-esse de licence universitaire, pour autant qu'un cours sur les variables aléatoires continues soit inclus dans le cursus.

Annexe 3

Ratio de Birge

Un des problèmes fondamentaux en métrologie consiste à savoir si plusieurs mesures sont bien compatibles entre elles. Une solution générale à ce problème a été proposée par Raymond T. Birge dans son article de 1932 [11, 16]. En reprenant les notations des deux annexes précédentes, il propose de comparer deux grandeurs :

- ◆ une estimation *statistique* de l'incertitude-type de la moyenne pondérée, calculé à l'aide d'un estimateur lui aussi pondéré (la pondération restant $w_i = \frac{1}{\sqrt{\langle X_i \rangle}} = \frac{1}{\sigma_{X_i}^2}$) ; il s'agit d'une estimation *externe* qui exploite la dispersion des différentes valeurs mesurées :

$$S_{\text{ext}} = \sqrt{\frac{1}{N-1} \frac{\sum_{i=1}^N w_i (X_i - \langle X \rangle)^2}{\sum_{i=1}^N w_i}} ;$$

- ◆ une estimation *a priori* de l'écart-type pour la moyenne pondérée, calculée à partir des différentes valeurs de $\sigma_{X_i} = \sqrt{\langle X_i \rangle}$; il s'agit d'une estimation *interne*, qui repose uniquement sur les différentes incertitudes-types rapportées :

$$S_{\text{int}} = \frac{1}{\sqrt{\sum_{i=0}^{N-1} \frac{1}{\sigma_{X_i}^2}}}.$$

Si les mesures sont compatibles entre elles, et donc compatibles avec le modèle qui les fonde, on s'attend à ce que ces deux grandeurs soient voisines, aux fluctuations statistiques près : l'estimation *interne* et *externe* doivent coïncider. Le ratio de Birge est le rapport de ces deux grandeurs, qui doit donc rester proche de l'unité lorsque les mesures sont compatibles.

Dans un cadre scientifique contemporain, on considère l'estimateur suivant, carré du ratio :

$$R^2 = \left(\frac{S_{\text{ext}}}{S_{\text{int}}} \right)^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} \left(\frac{X_i - \langle X \rangle}{\sigma_{X_i}} \right)^2.$$

Si une réalisation R_{exp}^2 de la variable aléatoire R^2 est suffisamment grande devant à 1, on est amené à rejeter la compatibilité des mesures. La dispersion des valeurs mesurées est plus grande que ce que prévoient les écarts-types. Il est bien sûr possible que les écarts-types aient sous-évalués. À l'inverse, si R_{exp}^2 est très inférieur à 1, on peut suspecter que les incertitudes ont été surévaluées. Une importante discussion métrologique à ce sujet est faite par B. N. Taylor [14].

Il s'agit ici d'une simplification de la pratique scientifique, nécessaire au niveau de la licence universitaire. Il arrive qu'on envisage plusieurs seuils avec divers plans d'action selon la valeur de R_{exp}^2 . Voir par exemple [15] pour la procédure suivie par le *Particle Data Group*, qui compile chaque année les propriétés des particules élémentaires et des interactions fondamentales. Il n'existe pas encore de procédure standardisée universellement acceptée, le sujet étant toujours débattu dans la littérature scientifique contemporaine [16–17].

Une fois la compatibilité rejetée, il convient d'examiner quelle mesure en particulier est en cause. On est alors amené à regarder lesquelles de ces mesures s'écartent notablement de la valeur moyenne, à l'instar de ce qu'on avait proposé au niveau lycée [2]. Il arrive parfois que seules les mesures extrêmes soient à la source de l'incompatibilité. On conclura alors que le modèle choisi est seulement valable dans un domaine restreint des paramètres.



Julien BROWAEYS

Maître de conférences

Laboratoire Matière et systèmes
complexes (MSC)

Université de Paris

Paris