

Mesures, erreurs et incertitudes en physique-chimie

René Moreau,

Inspecteur général de l'éducation nationale

Introduction

Première réflexion sur le contenu du document

Ce document peut apparaître à certains comme ne contenant que des notions triviales, tandis qu'il donnera à d'autres l'impression de traiter de sujets qu'il serait impossible de présenter aux élèves.

Aux uns comme aux autres, par avance, je conseillerai de nuancer leur opinion. Aux premiers, je demande de faire l'effort de se montrer attentifs bien qu'ils connaissent les notions théoriques qui seront mobilisées dans cet exposé : il m'est en effet arrivé bien souvent d'entendre un professeur m'affirmer qu'il attachait toujours une grande importance aux notions qui font l'objet de ce travail, alors que les documents qu'il me remettait (textes d'exercices ou de TP donnés à ses élèves, de problèmes construits pour eux) fourmillaient d'erreurs et d'exemples de non maîtrise du sujet.

Reconnaissons à l'intention des seconds que le sujet est assez complexe et qu'il comporte de nombreuses facettes. Sachons aussi qu'on ne peut attendre que toutes les notions qui fondent son étude soient maîtrisées par les élèves pour commencer à sensibiliser ces derniers aux problèmes des mesures, si bien que nombre de savoir-faire peuvent leur apparaître comme des recettes (ce qui choque plus les professeurs que les élèves). Mais il faut que ces réticents sachent que la majorité des notions nécessaires au traitement des mesures effectuées en classe sont assez intuitives, qu'aucune ne heurte le bon sens, si bien que les élèves admettent fort bien que certaines règles leur soient données sans démonstration et cela dès le collège. De plus, à l'heure actuelle, l'utilisation de logiciels commodes et clairs facilite le travail des professeurs de lycée. Le traitement raisonné des mesures des élèves donne aux professeurs tant d'occasions de donner plus de sens à leurs cours, de mettre leurs élèves sur le bon chemin quant à l'interprétation des résultats, bref d'augmenter sensiblement l'intérêt des séances d'enseignement qu'ils dispensent, que cela vaut largement la peine d'essayer. S'ils y croient vraiment, les professeurs percevront en plus une sorte de prime : par des observations qui sont de la même nature que celles d'un chercheur, mais qui sont appliquées à nos travaux pratiques, le traitement des mesures des élèves permet souvent de faire la lumière sur tel ou tel phénomène qui *a priori*, ne nous apparaissait pas comme important mais dont une analyse plus fine révèle l'influence. L'intérêt de notre métier s'en trouve alors singulièrement accru.

Afin que ces dernières remarques n'apparaissent pas comme trop théoriques, en m'appuyant sur le travail de dizaines de professeurs avec lesquels, en un quart de siècle, j'ai pu échanger des idées à propos des mesures de leurs élèves, je citerai la mise en évidence de nombreuses erreurs systématiques. En 1977, par exemple, en seconde C, au lycée d'Agen, dans l'expérience de Gay-Lussac conduisant à l'expression $p = p_0 (1 + \alpha t)$ pour un volume d'air maintenu constant et dont la température t était mesurée en degrés Celsius, la mise en commun du travail de deux classes montra que le coefficient α n'était pas tout à fait égal à $\frac{1}{273,15}$ et l'analyse attentive des bouchons montra qu'ils fuyaient légèrement. Un peu plus tard, au lycée Montaigne de Bordeaux, en terminale, l'étude de la chute libre des corps, à l'aide d'un dispositif qui mettait en œuvre une bille en acier, parce qu'elle conduisait à une valeur de g un peu trop faible, montra que la bille ne devenait relativement libre qu'après s'être suffisamment éloignée de l'électroaimant qui l'attirait encore un peu, dans le premier centimètre de sa chute. Plus récemment, au lycée David d'Angers, le décalage persistant entre résultat attendu d'un dosage et le résultat moyen

obtenu, a montré l'influence néfaste d'une balance de précision (mais qui n'avait pas été ré-étalonnée) à partir de laquelle les produits chimiques étaient pesés par le technicien de laboratoire.

Dans la belle expérience que réalisent les élèves de première S afin de mesurer le volume molaire des gaz parfaits, l'influence de la pression de vapeur saturante de l'eau sur le volume d'hydrogène que les élèves recueillent (dont la non prise en compte provoque une erreur systématique de l'ordre de 2,5 %) est apparue évidente. Cette même expérience (mais cela aurait tout aussi pu se produire à propos de la mesure de la vitesse du son dans l'air) a également amené les professeurs à réaliser que, lorsqu'on téléphone à un laboratoire de météorologie, pour connaître la pression atmosphérique locale, celui-ci ne vous donne pas la pression dans la ville où vous vivez, mais la pression corrigée, ramenée au niveau de la mer, car c'est elle qui est pertinente pour la prévision des phénomènes météorologiques. Or la pression atmosphérique chute de 10 hPa (soit 1 % de sa valeur) lorsque l'altitude s'élève de 85 m : de nombreuses mesures d'élèves, si elles sont convenablement exploitées en classe, sont affectées par cette correction et rendent nécessaire la prise en compte de l'altitude de la salle de classe.

Les buts que se propose d'atteindre ce document

Les mesures sont entachées d'erreurs

Nous savons tous qu'une mesure physique, aussi précise soit-elle, comporte toujours une part d'incertitude et probablement, une erreur, même minime. Mais si nous ne montrons jamais que, malgré nos efforts, de telles erreurs entachent nos mesures et si nous n'utilisons jamais le vocable d'incertitude, pour montrer que nous sommes conscients de ces erreurs, alors, sur ce point, notre enseignement est muet.

Les erreurs susceptibles d'entacher la mesure d'une grandeur sont multiples et la recherche de leurs causes peut être passionnante. Citons brièvement l'environnement de la grandeur mesurée, qui peut être influencée par d'autres grandeurs, l'équipement de mesure dont on dispose, la qualification des opérateurs effectuant la mesure ou la méthode utilisée.

En classe, la répétition de certaines mesures, selon un même mode opératoire, par des groupes d'élèves de même niveau, utilisant du matériel comparable, peut rapidement donner une idée exploitable de la répartition des erreurs, à l'exception des erreurs systématiques. Mais la comparaison des mesures obtenues par deux méthodes différentes (focométrie, dosages), ou la mesure de grandeurs déterminées avec précision par ailleurs, permet aussi de débusquer certaines erreurs systématiques.

Présentation des résultats de mesure

De manière idéale, une grandeur ayant pour valeur exacte (inconnue) la quantité A , le résultat de la mesure de A devrait être donné sous la forme $a = \hat{a} \pm \Delta a$: \hat{a} est l'estimateur de la valeur exacte A (par exemple la moyenne de quelques mesures) et Δa l'incertitude sur la mesure de A , telle que la probabilité P pour que l'intervalle $(\hat{a} - \Delta a, \hat{a} + \Delta a)$ contienne A soit assez élevée (par exemple $P = 0,95$).

Cette écriture suppose que \hat{a} ne coïncide pas exactement avec A : $\varepsilon_a = \hat{a} - A$ est l'erreur commise sur la mesure de A . On présentait autrefois Δa , l'incertitude absolue sur A , comme un majorant de $|\varepsilon_a|$: tout le calcul d'incertitudes classique en découlait, notamment les théorèmes sur la composition des incertitudes absolues ou celle des incertitudes relatives du type $\Delta a/\hat{a}$. Ces théorèmes conduisaient à des impasses : on "montrait" ainsi que la moyenne m de n mesures indépendantes $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$, n'était pas plus précise que l'une quelconque des mesures, alors que, dans une classe, notamment, chacun peut (et doit) constater le contraire.

En fait, chaque opérateur, lorsqu'il s'inscrit dans la logique visant à déterminer un *majorant* Δa de $|\varepsilon_a|$, minimise inconsciemment l'incertitude Δa , comme le montrent par exemple les mesures historiques successives de la vitesse c de la lumière par des savants, qui, presque tous, ont proposé des majorants de $|\varepsilon_c|$ qui se sont révélés trop faibles.

De nombreux constructeurs d'appareils de mesure utilisés dans les établissements scolaires procèdent de même. Au bout d'un an de fonctionnement, par exemple, il n'est pas rare de trouver pour ces appareils des erreurs trois fois plus grandes que le "majorant" indiqué par la notice ! Il est donc beaucoup plus

réaliste de considérer que l'intervalle $(\hat{a} - \Delta a, \hat{a} + \Delta a)$ est un *intervalle de confiance*, la probabilité P qui lui est associée ($P = 0,95 = 95 \%$, ou $P = 0,99$) étant le *niveau de confiance* de l'intervalle choisi. Bien sûr, l'incertitude Δa est d'autant plus grande que le niveau de confiance P se rapproche de l'unité.

Ce point de vue, qui considère les différentes mesures indépendantes a_i de A comme les valeurs successives prises par la *variable aléatoire* a , se heurte cependant à une importante difficulté : pour affecter à une incertitude Δa un niveau de confiance P donné, il faut connaître la loi de répartition des mesures a_i autour de A , encore appelée densité de probabilité de a .

On suppose en général qu'il s'agit d'une répartition *gaussienne*, que chacun connaît au moins approximativement et dont on peut rappeler qu'elle est symétrique par rapport à sa valeur moyenne X et pour laquelle les faibles écarts à la moyenne sont plus probables que les forts.

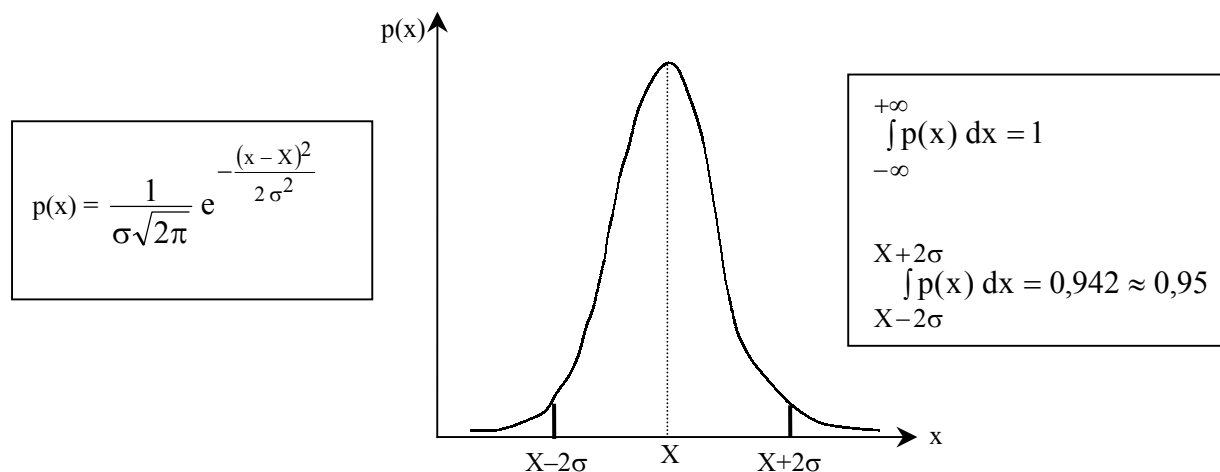


Figure 1

Cette loi est si répandue qu'on la dénomme la loi *normale* de Gauss. Cette supposition est en général légitime (il suffit par exemple, ce sont les "conditions de Borel", que les causes d'erreurs soient multiples et d'importance comparable) ; cette loi de répartition a d'ailleurs été prouvée dans la quasi-totalité des cas où tel ou tel d'entre nous a pris le temps de regrouper un millier de mesures *indépendantes* et *de poids égaux* d'une même grandeur. En général, faute de temps, l'hypothèse gaussienne ne peut ni être prouvée, ni réfutée. Cependant, le logiciel « Incertitudes de mesure »¹, mis au point par le CNDP sous ma direction (bénévole), permet instantanément de savoir si l'échantillon de mesures dont on dispose peut ou non être considéré comme tiré d'une population gaussienne. Il informe également l'opérateur de l'existence d'une ou plusieurs mesures aberrantes et, si l'opérateur accepte d'enlever ces valeurs, le logiciel calcule automatiquement l'incertitude Δa qui correspond à ces mesures en utilisant pour cela des critères sur lesquels je reviendrai et qui peuvent être admis par tous.

Il est bon de savoir qu'en matière d'Incertitudes de mesure, les méthodes statistiques normalisées se sont définitivement imposées au début des années 90, d'abord dans les laboratoires de métrologie (en particulier au Bureau national de métrologie), puis dans les entreprises, avec l'installation des normes d'assurance qualité ISO 9000.

Certains laboratoires de métrologie accompagnent seulement un résultat de mesure, considéré comme une valeur de variable aléatoire, de l'écart-type correspondant. Cet écart-type est baptisé "incertitude type", tandis qu'on appelle "incertitude élargie" la quantité 2σ : la valeur $k = 2$ qui a été retenue au niveau européen comme la valeur standard à prendre en compte, correspond, lorsque la répartition des mesures est gaussienne, à un niveau de confiance voisin de 95 % généralement considéré comme suffisant (c'est $1,96\sigma$ qui correspond exactement au niveau de confiance 95 %) : nous sommes très près des idées qui seront exposées plus loin comme pouvant être appliquées en classe.

¹ Cndp, coll. « Simulation et analyse de données ». Référence : 755 03624.

On pourrait s'en tenir à la seule notion d'écart-type. Elle permet notamment un nouveau type de "calculs d'incertitude", plus réaliste que l'ancien, fondé sur les seules propriétés de la variance des fonctions aléatoires et indépendant des lois de probabilités. Mais ces calculs, qui supposent maîtrisé le calcul différentiel, ne peuvent être menés qu'au niveau de l'enseignement supérieur. Ils ne sont pas envisageables au lycée où le problème de l'estimation d'une incertitude *raisonnable* ne peut être résolu que par des méthodes *directes* et *globales*. C'est d'ailleurs largement suffisant.

Pour initier les élèves aux méthodes statistiques utilisées par la mesure moderne, mais aussi dans d'autres types d'investigations, fixons-nous donc pour objectif de faire déterminer de temps en temps par les élèves de première ou de terminale S, à partir de leurs propres pratiques expérimentales et en regroupant seulement les mesures indépendantes de leur groupe de TP, sans lourds calculs mathématiques, un intervalle de confiance au niveau de confiance 95 %. Cet objectif à long terme peut être préparé dans les classes précédentes et même dès le collège.

La question des chiffres significatifs

Les exercices proposés aux élèves comportent des données exprimées avec un certain nombre de chiffres significatifs. Les réponses numériques à ces exercices doivent être cohérentes avec les données. Rechercher cette cohérence, la justifier, est une activité qui s'apparente à la réflexion sur les incertitudes de mesure. On peut donc, par ce biais, sensibiliser très tôt les élèves à ces questions. Dans ce seul domaine, d'immenses progrès restent à accomplir, à tous les niveaux, c'est-à-dire du collège à l'agrégation, car il n'est pas rare de trouver dans une épreuve d'agrégation des données exprimées par exemple avec un seul chiffre significatif alors que les examinateurs attendent des résultats exprimés avec trois chiffres significatifs. On aura par exemple un condensateur de capacité $C = 1 \text{ nF}$ et il faut donner la valeur de l'inductance L qui, associée à C , provoquera une résonance électrique à $f_0 = 120 \text{ kHz}$. L'application de la formule $LC\omega^2 = 1$ conduit les candidats à écrire $L = 1,76 \text{ mH}$ ou même $1,759 \text{ mH}$. Alors que, si l'auteur du sujet avait pensé au problème des chiffres significatifs, il aurait donné, par exemple, $C = 1,0 \text{ nF}$, $f_0 = 1,2 \times 10^5 \text{ Hz}$ et il était en droit de s'attendre à ce qu'on lui réponde $L = 1,8 \text{ mH}$.

Ou $C = 1,00 \text{ nF}$ et $f_0 = 120 \text{ kHz}$, ce qui entraîne bien $L = 1,76 \text{ mH}$.

Par ailleurs, l'écriture $\hat{a} \pm \Delta a$ suppose que l'on choisisse les chiffres significatifs qui expriment \hat{a} et Δa . En nous fondant sur des considérations statistiques, nous proposerons, en annexe 5, des critères quantitatifs permettant d'effectuer logiquement ce choix. Avec ces critères, on remarque que ceux des professeurs qui s'intéressent au sujet ont parfois tendance à enlever trop de chiffres significatifs, aussi bien pour exprimer l'estimateur \hat{a} que l'incertitude Δa . Cependant, ces critères sont un peu difficiles à mettre en œuvre ; il faut donc appliquer des règles d'emploi plus simple et qui font à peu près l'unanimité (et surtout, il ne faut pas adopter une attitude rigide pour l'application de ces règles qui ne correspondent que de manière approchée à l'application de critères scientifiques). Pour cela, les consignes données il y a bien longtemps par le grand physicien que fut G. Bruhat, dans son traité de Mécanique (page 265 de la sixième édition, chez Masson), restent tout à fait d'actualité : après avoir trouvé, par exemple, pour l'indice d'un prisme $n = 1,50944 \pm 0,00039$, il recommandait d'écrire : $n = 1,5094 \pm 0,0004$.

"L'incertitude étant de plusieurs unités de la quatrième décimale, il est inutile, disait-il, de conserver la cinquième décimale, mais il faut garder la quatrième car en la supprimant on introduirait une erreur d'écriture parfaitement inutile, comparable à l'erreur expérimentale".

De la même manière, partant par exemple d'une mesure de puissance telle que $P = (3420,5 \pm 25,67) \text{ W}$, nous verrons par la suite que l'écriture définitive du résultat devrait être : $P = (3,420 \pm 0,026) \text{ kW}$. En gros, l'incertitude Δa ne doit jamais être exprimée avec plus de deux chiffres significatifs et, pour l'estimateur \hat{a} , on conserve tous les chiffres significatifs qui sont affectés par Δa .

Quand et comment introduire ces notions dans les classes ?

Depuis vingt ans, mis à part quelques cas isolés, ces notions ne sont presque pas abordées. Cela conduit à des aberrations, car les jeunes gens qui sortent du lycée, y compris ceux qui, en première et en

terminale ont choisi de participer aux Olympiades de la Physique, n'ont aucune idée de ce que peut représenter une incertitude de mesure, qu'elle soit considérée comme majorant ou affectée d'une probabilité limitée ; l'incohérence s'installe même en ce qui concerne les chiffres significatifs des données et des résultats.

Nous ne pouvons continuer à ignorer le sujet et il faut donc progressivement tenter de "remonter la pente". Le mieux serait d'aborder ces questions dès le collège. Évidemment, c'est de manière prudente qu'il faut agir à ce niveau : il ne peut y être question, ni d'intervalle de confiance, ni d'écart-type. Mais on peut montrer que la précision d'une mesure est limitée et adopter des règles simples sur la question des chiffres significatifs. La sensibilisation précoce des jeunes élèves aux problèmes de la mesure, lorsqu'ils manipulent, permettrait de leur expliquer certaines contradictions apparentes et de répondre à leurs questions. En revanche, à aucun niveau de l'enseignement secondaire, il ne faut utiliser le modèle handicapant du majorant des erreurs qui, en conduisant aux anciennes règles du calcul d'incertitude, deviendrait plus tard gênant pour eux (pas plus qu'aucun modèle théorique de "propagation des erreurs").

Nous allons donc tenter de proposer une introduction progressive des notions dont il vient d'être question, introduction ni dogmatique ni abstraite, mais réalisée à partir des activités expérimentales réalisées en classe, soit dans les expériences de cours à exploitation collective, soit à l'occasion des travaux pratiques des élèves.

Nous commencerons donc par les classes du collège auxquelles nous associons la classe de 2^{de} indifférenciée. La seconde est en effet désormais une classe où les notions de physique sont rares, comme les occasions de pratiquer de véritables mesures.

Les classes de 1^{ère} et de terminale S ou industrielles constituent un deuxième stade, très différent, où les élèves ont en principe choisi de faire des sciences.

Des exemples d'applications effectivement réalisés dans les classes illustreront notre propos.

Ce qui peut être introduit au collège et en seconde

Généralités

Au collège, en cinquième et en quatrième, les élèves mesurent des températures pour repérer les changements d'état, ils manipulent ampèremètres et voltmètres ; ils peuvent mettre en évidence la variation de volume de l'eau au cours de la fusion de la glace... En troisième, on mesure des distances et des temps, pour accéder à la notion de vitesse moyenne, tandis que la relation entre masse et poids invite à mesurer ces deux types de grandeurs. Pour une résistance, l'ohmmètre permet de comparer son indication au quotient U/I des indications de deux appareils déjà connus. En chimie, enfin, l'usage du pH-mètre n'est pas exclu... On pourrait multiplier les exemples qui posent le problème de la mesure et donc celui des chiffres significatifs, car, après un mesurage, effectué par les élèves ou le professeur, on doit en effet afficher le résultat, la mesure, et utiliser pour cela une unité et un certain nombre de chiffres significatifs.

Très tôt, en quatrième par exemple, il faut montrer aux élèves que les indications fournies par les appareils de mesure sont entachées de petites erreurs et que, de ce fait, et, parfois aussi, du fait de l'expérimentateur lui-même, il n'y a pas tout à fait coïncidence entre la valeur exacte A d'une grandeur (parfois connue par ailleurs) et la mesure \hat{a} correspondante. On peut, de temps en temps, proposer aux élèves d'évaluer expérimentalement l'ordre de grandeur de l'erreur commise sur une mesure, autrement dit, s'intéresser à l'incertitude qui lui est liée.

En troisième, on peut remarquer que de nombreuses grandeurs mesurées dépendent légèrement de facteurs physiques (température, pression...) susceptibles d'évoluer en fonction du temps de manière incontrôlée. D'autres, comme les dimensions d'une table rectangulaire en bois, le diamètre d'une sphère imparfaite, ne sont définies qu'avec une certaine précision.

Ces réalités n'ont pas à être cachées (il ne faut donc pas, pour "vérifier" la loi des courants dérivés $i = i_1 + i_2$, choisir, comme on le fait parfois, des courants de 100 à 200 mA et utiliser, pour les mesurer, les calibres 10 A des trois ampèremètres utilisés !).

Les observations liées au mesurage conduiront ainsi professeurs et élèves à présenter la plupart des mesures et calculs effectués en classe, avec, simplement, deux ou trois chiffres significatifs et l'emploi correct des multiples et sous-multiples des unités légales ($m = 12,3 \text{ g}$; $V = 83 \text{ cm}^3$; $P = 6,5 \text{ N}$; $g = 9,8 \text{ N/kg}$; $I = 24,5 \text{ mA}$; $U = 4,5 \text{ V}$; $R = 47,3 \text{ k}\Omega$; $\theta = 99,5^\circ\text{C}$; $P = 105 \text{ W} \dots$).

Certains résultats, ayant fait l'objet d'une réflexion plus approfondie, pourront être donnés sous la forme d'un intervalle centré, par exemple : $g = (g_0 \pm \Delta g) \text{ N/kg}$. Dans ce cas, l'intervalle $(g_0 - \Delta g, g_0 + \Delta g)$ sera présenté comme ayant "de fortes chances" de contenir la "vraie valeur" de g . Il s'agit de sensibiliser à la notion d'incertitude, et non de faire œuvre définitive. Ne pas dire de choses fausses, mais ne pas chercher non plus, à établir du premier coup la "doctrine" achevée, car c'est absolument impossible.

Les chiffres significatifs au collège

Il s'agit d'un problème difficile : la manière dont on s'exprime dans la vie courante n'en favorise pas la solution qui, de toute manière, n'est pas simple. En ville, quand une personne dit qu'elle a payé sa voiture neuve 110 000 F, cela signifie généralement qu'elle ne considère comme significatifs que les seuls deux premiers chiffres non nuls de cette somme alors que le prix réel, avec le plein d'essence, la carte grise et les plaques minéralogiques, était peut être 113 875,23 F ou, aussi bien, 108 256,89 F. Si aucun des quatre zéros n'est considéré comme significatif, la "bonne" manière de procéder, indiquant que le prix P payé est compris entre 105 000 F et 115 000 F (prix à 5 kF près en valeur absolue, ou encore à 5 % près en valeur relative), consisterait à écrire $P = 1,1 \times 10^5 \text{ F}$ ce qui apparaîtrait comme pédant. De même, si l'on désire indiquer que le prix est compris entre 109 500 F et 110 500 F (prix à 500 F près en valeur absolue ou à 0,5 % près en valeur relative), le premier zéro est un chiffre significatif, il faudrait dire que l'on a payé sa voiture 110 kF.

En mathématiques, écrire $m = 11 597 \text{ g}$, signifie que seul le dernier chiffre, 7, est incertain. On a donc $11 596,5 \text{ g} \leq m \leq 11 597,5 \text{ g}$. En effet, si le résultat trouvé est 11 597,7 g, par exemple, alors la masse m , arrondie au gramme, devient 11 598 g. En physique, en l'absence d'indication explicite sur l'incertitude attachée à m , on admet souvent que celle-ci est égale à une demi unité du dernier chiffre exprimé (P. Fleury et J.-P. Mathieu, Mécanique physique, 4ème édition, 1965, page 42). Les écritures $m = 11 597 \text{ g}$ ou $m = (11 597 \pm 0,5) \text{ g}$ sont donc équivalentes. En revanche, si l'on désire indiquer que l'incertitude Δm sur m est de 1 g, par exemple, au sens où l'intervalle (11 596 g, 11 598 g) a de fortes chances de contenir la vraie valeur de m , alors il faut écrire $m = (11 597 \pm 1) \text{ g}$.

Dans ce cas, on a donc mesuré une masse de plus de 10 kg à 1 g près, soit avec une précision relative $\frac{\Delta m}{m} = 1 \times 10^{-4}$. Si des précisions de cet ordre sont très facilement réalisées en métrologie, elles sont rares dans la vie courante.

Au collège, la précision relative des mesures est couramment de l'ordre de 10^{-2} , soit 1 % en pourcentage. Cela entraîne en général l'écriture des mesures avec deux ou trois chiffres significatifs, par exemple :

$$m_1 = 11,6 \text{ kg} ; m_2 = 8,7 \text{ g} ; m_3 = 0,79 \text{ kg} ; m_4 = 228 \text{ g}.$$

Dans certains cas (appareils neufs ou de très bonne qualité ayant fait l'objet d'un suivi de maintenance), la précision relative atteint 10^{-3} . La mesure correspondante s'écrit alors avec 3 chiffres significatifs ($m_3 = 795 \text{ g}$), ou quatre ($U = 10,56 \text{ V}$).

Dans les calculs où il n'y a que des multiplications ou des divisions de grandeurs, du type $z = x \text{ y}$ ou $z = \frac{x}{y}$, la précision relative $\frac{\Delta z}{z}$ du résultat final, z , ne peut être meilleure que celle de la grandeur x ou y la moins précise : le rapport $\frac{\Delta z}{z}$ est donc supérieur au plus grand des rapports $\frac{\Delta x}{x}$ ou $\frac{\Delta y}{y}$.

Dans l'hypothèse où Δx , Δy et Δz sont, comme en mathématiques, des majorants des erreurs commises sur x , y et z , ce résultat peut, dès le collège, être "justifié" sur des exemples numériques. Toujours dans cette hypothèse, il peut être établi rigoureusement en terminale S. Cependant, ce point de vue est incompatible avec l'objectif recherché qui consiste seulement à associer aux intervalles $[x - \Delta x, x + \Delta x]$, $[y - \Delta y, y + \Delta y]$, $[z - \Delta z, z + \Delta z]$ une probabilité (par exemple $P \approx 0,95$) de contenir respectivement les valeurs exactes X , Y et Z . Dans ce cas, ce n'est qu'en classe post-baccalauréat qu'on peut démontrer le théorème précédent, mais nous verrons qu'on peut, à titre d'exercice, le faire toucher du doigt en première.

Supposons ainsi que l'on s'intéresse, à la vitesse moyenne v_{moy} sur 1000 m d'une voiture qui, arrêtée au départ, parcourt cette distance en 31,4 s ; ceci afin, par exemple, de la comparer à la vitesse maximale annoncée. En l'absence de précision supplémentaire sur ces mesures, nous remarquons que la distance L est donnée avec une précision relative de $0,5 \times 10^{-3}$ et la durée t avec une précision relative de $1,6 \times 10^{-3}$. De ce fait, on ne peut donner la vitesse moyenne $v_{\text{moy}} = L/t$ avec une précision meilleure que $1,6 \times 10^{-3}$. La bonne manière d'exprimer le résultat est donc $v_{\text{moy}} = 115 \text{ km/h}$. La vitesse 114,6 km/h est trop précise (précision relative de $0,44 \times 10^{-3}$), celle de $1,1 \times 10^2 \text{ km/h}$ (précision relative 45×10^{-3}) ne l'est pas assez.

On ne commet pas d'erreur grossière en remplaçant la règle ci-dessus par celle qui consiste à donner le résultat de la multiplication ou de la division avec un nombre de chiffres significatifs égal à celui de la donnée la moins précise.

Prenons un autre exemple : la mesure "à chaud" d'une résistance a donné $R = 47,2 \Omega$; cette résistance est parcourue par un courant d'intensité $I = 68 \text{ mA}$. On demande de calculer la tension U aux bornes de R .

Observons qu'en l'absence de précision supplémentaire, la donnée $R = 47,2 \Omega$ signifie que l'incertitude absolue sur R est de $0,05 \Omega$, soit une incertitude relative $\Delta R/R = 1 \times 10^{-3}$. De même, et avec la même restriction, on fait l'hypothèse que $\Delta I/I = 7,3 \times 10^{-3}$. Dans ces conditions, alors que le produit RI est égal à $3,2096 \text{ V}$, la précision relative $\Delta U/U$ de la tension $U = RI$ ne peut être meilleure que $7,3 \times 10^{-3}$. L'incertitude ΔU est donc supérieure à 23 mV et l'on ne peut donc écrire $U = 3,21 \text{ V}$, ce qui sous-entend que cette incertitude est égale à 5 mV . Le résultat cohérent avec les données est donc $U = 3,2 \text{ V}$.

Le marquage des résistances radio met bien l'accent sur les chiffres significatifs : on ne peut que recommander son utilisation.

Pour les résistances habituelles (de tolérance 5 % ou 10 %), deux cercles de couleur indiquent les chiffres significatifs, un troisième le multiplicateur et le quatrième la tolérance (ou précision relative). Ainsi, une résistance comportant un cercle bleu (6), un gris (8), un jaune (4) et un doré a pour valeur nominale $68 \times 10^4 \Omega$, soit $0,68 \text{ M}\Omega$ à 5 % près, et non, comme on le voit parfois écrit au tableau, $680\,000 \Omega$, ni même $680 \text{ k}\Omega$. Le fabricant garantit en effet que, dans les limites normales prévues pour son fonctionnement (limites portant sur la température ambiante et la puissance dissipée), sa valeur est comprise entre $646 \text{ k}\Omega$ et $714 \text{ k}\Omega$ et non entre $679\,999,5 \Omega$ et $680\,000,5 \Omega$, ni entre $679,5 \text{ k}\Omega$ et $680,5 \text{ k}\Omega$ ce qui supposerait encore une précision meilleure que $10^{-3} \Omega$.

L'emploi des puissances de 10 pour l'écriture des résultats, ou de "la notation ingénieur" recommandée dans le programme de mathématiques de la classe de quatrième, ou, ce qui revient pratiquement au même, l'emploi courant des multiples et sous-multiples des unités légales, est indispensable pour maîtriser la question des chiffres significatifs. Nous devons, en physique-chimie, habituer les élèves, dès le collège, à ne pas considérer comme totalement équivalentes des mesures différant les unes des autres par le nombre de chiffres significatifs. En particulier, les exercices de "conversions d'unités", nombreux, à juste titre, dans les manuels, devraient veiller au respect de ces nombres : si la capacité V d'une citerne est $2,75 \text{ m}^3$, elle peut certes s'exprimer en mL, mais en écrivant $V = 2,75 \times 10^6 \text{ mL}$ et non $V = 2\,750\,000 \text{ mL}$, car cette dernière écriture, qui sous-entend que la capacité est connue au millilitre près, n'est pas équivalente à l'information initiale. De même, la consommation journalière en charbon d'une chaudière d'immeuble est 15 tonnes (ou $15,0 \text{ t}$) plutôt que $15\,000\,000 \text{ grammes}$!

Donnons encore l'exemple de la distance moyenne de la Terre autour du Soleil (moyenne temporelle), il s'agit de "l'unité astronomique" très proche de la longueur du demi grand axe de l'orbite de la Terre autour du Soleil :

$1 \text{ UA} = 1,495\,978\,70 \times 10^{11} \text{ m}$. On prend souvent, à juste titre, pour valeur arrondie de cette distance,

cent cinquante millions de kilomètres. Mais cela doit s'écrire :

$1 \text{ UA} = 1,50 \times 10^8 \text{ km}$, ou $1 \text{ UA} = 1,50 \times 10^{11} \text{ m}$, et non $150\,000\,000 \text{ km}$. Cette dernière écriture, en effet, semble donner la valeur de l'unité astronomique arrondie au kilomètre, alors que la valeur retenue par les astronomes en diffère de $402\,130 \text{ km}$.

La grandeur mesurée varie parfois quelque peu, de manière non maîtrisée

Prenons le cas de la résistance d'un dipôle ohmique classique de valeur nominale 47Ω à 5 % près (résistance "radio" portant les cercles jaune, violet, noir et doré). Cette résistance est comprise entre $44,5 \Omega$ et $49,5 \Omega$, mais, en général, elle est plus proche de 47Ω que ne le laisse penser la largeur de cet intervalle. Mesurée à l'aide d'un ohmmètre fournissant 4 chiffres significatifs, on obtiendra par exemple $R = 46,93 \Omega$ à la température de la salle. Supposons qu'il s'agisse d'une résistance destinée à être insérée dans un montage et capable de dissiper au maximum la puissance $0,5 \text{ W}$ ce qui advient lorsque la tension à ses bornes atteint $4,8 \text{ V}$. Quand on applique effectivement cette tension, la température moyenne de la résistance radio augmente souvent d'une quinzaine de degrés et, lorsque l'équilibre thermique est atteint, la nouvelle valeur de R , mesurée avec le même ohmmètre devient par exemple $R' = 46,73 \Omega$. D'ailleurs, en serrant simplement la résistance entre le pouce et l'index, la valeur affichée passe alors, par exemple, de $46,93 \Omega$ à $46,82 \Omega$ en une ou deux minutes si la température de la résistance varie d'une dizaine de degrés au cours de cette opération (ce qui peut être le cas si la température de la salle est voisine de 18°C).

R varie donc avec la température et l'ensemble de ses valeurs, correspondant à un usage "normal", à température ambiante constante, occupe une plage de largeur voisine de $0,2 \Omega$.

On en tire tout de suite une première conséquence : le quatrième chiffre significatif donné par l'appareil, correspondant à $10 \text{ m}\Omega$, n'a pas à être conservé puisqu'une variation de température de quelques degrés, couramment subie par la résistance, entraîne une variation de R dix fois supérieure ($0,1 \Omega$) : on peut retenir $R = 46,9 \Omega$ comme valeur de la résistance à la température de la salle et $R = 46,7 \Omega$ lorsque la résistance s'est échauffée en dissipant sa puissance maximale (c'est la valeur "à chaud"). Si la puissance dissipée par la résistance est susceptible de varier entre 0 et $0,5 \text{ W}$, on estimera mieux sa valeur moyenne R en prenant la demi-somme de ces deux valeurs, soit $R = 46,8 \Omega$.

On peut minimiser de telles variations en choisissant bien son matériel. Dans le cas particulier précédent, (résistance de 47Ω soumise à des tensions comprises entre 0 et 5 V), supposons que l'on utilise une résistance métallique conçue pour pouvoir dissiper 50 W . Celle-ci est construite pour que sa variation n'excède pas $0,2 \Omega$ lorsqu'elle dissipe effectivement 50 W (sa température augmente alors d'une centaine de degrés Celsius). Ainsi, pour de faibles variations de température (5 ou 6 C), correspondant à des courants inférieurs à $0,3 \text{ A}$, la valeur d'une telle résistance métallique est-elle bien définie à $10 \text{ m}\Omega$ près.

Nous venons d'aborder le cas d'une grandeur qui varie quelque peu en fonction de la température. On peut aisément observer que d'autres grandeurs, que l'on souhaiterait constantes, dépendent légèrement de paramètres extérieurs :

La température d'un liquide que l'on chauffe par le fond du contenant, ou qui se refroidit par sa surface libre, ou d'un liquide qui bout, n'est pas parfaitement homogène : au collège, on peut espérer déterminer une telle température au demi degré Celsius près, mais non au dixième.

L'intensité du courant qui circule dans un circuit varie toujours un peu avec le temps, parce que les résistances du circuit et le générateur s'échauffent (ou la pile s'use) : on constate facilement qu'un courant réglé à la valeur $i = 100,0 \text{ mA}$, dérive par exemple jusqu'à $99,5 \text{ mA}$ en quelques minutes.

La masse d'un liquide varie, lorsqu'il se trouve dans un récipient non fermé, parce que le liquide s'évapore plus ou moins rapidement.

Les variations au cours du temps de la valeur efficace U de la tension du secteur, voisine de 230 V , sont en général largement supérieures à 1 V .

Le pH d'une solution aqueuse faiblement concentrée, laissée à l'air libre dans une salle de classe occupée, peut varier avec le temps car du dioxyde de carbone s'y dissout, etc.

Les erreurs des appareils de mesure et des manipulateurs

Lorsque l'on désire vérifier si les indications de trois ampèremètres différents, utilisés pour mesurer I , I_1 et I_2 , sont en conformité avec la loi $I = I_1 + I_2$ des courants dérivés, il arrive couramment, avec les appareils numériques actuels, que les mesures respectives \hat{i} , \hat{i}_1 et \hat{i}_2 des intensités I , I_1 et I_2 ne satisfassent pas rigoureusement à l'égalité $\hat{i} = \hat{i}_1 + \hat{i}_2$. Il ne faut bien évidemment pas en laisser déduire que la loi n'est qu'approchée, mais traiter la difficulté. Supposons par exemple qu'on utilise des appareils assez récents dont la notice indique que l'incertitude Δi attachée à chaque mesure est telle que $\Delta i = 1\% \text{ VL} + 1 \text{ UR}$: VL est la valeur lue, tandis que le sigle 1 UR (une unité de représentation), signifie que l'on doit ajouter la valeur correspondant à une unité du dernier chiffre affiché. Par exemple, si $\hat{i} = 187,2 \text{ mA}$, alors $1\% \text{ VL} = 1,9 \text{ mA}$ et $1 \text{ UR} = 0,1 \text{ mA}$ soit $\Delta i = 2,0 \text{ mA}$.

Les mesures $\hat{i}_1 = 50,1 \text{ mA}$, $\hat{i}_2 = 138,6 \text{ mA}$, et $\hat{i} = 187,2 \text{ mA}$, par exemple, ne mettent pas en cause l'égalité $I = I_1 + I_2$, et pourtant $\hat{i}_1 + \hat{i}_2 = 188,7 \text{ mA} \neq \hat{i} = 187,2 \text{ mA}$.

Elles sont même parfaitement compatibles avec elle puisque la différence $(\hat{i}_1 + \hat{i}_2) - \hat{i} = 1,5 \text{ mA}$ est inférieure à la seule incertitude $\Delta i = 2,0 \text{ mA}$.

Remarque 1 : Les mesures \hat{i} , \hat{i}_1 et \hat{i}_2 sont respectivement entachées des erreurs $\varepsilon_i = \hat{i} - I$, $\varepsilon_{i_1} = \hat{i}_1 - I_1$, et $\varepsilon_{i_2} = \hat{i}_2 - I_2$. Si les incertitudes Δi , Δi_1 et Δi_2 indiquées par le constructeur étaient effectivement les majorants respectifs de $|\varepsilon_i|$, $|\varepsilon_{i_1}|$ et $|\varepsilon_{i_2}|$, alors la somme $(\Delta i + \Delta i_1 + \Delta i_2)$, soit $(2,0 + 0,6 + 1,5) \text{ mA} = 4,1 \text{ mA}$, serait bien un majorant de la différence $(\hat{i}_1 + \hat{i}_2 - \hat{i})$. Mais ce n'est généralement pas le cas : l'expérience montre en effet que les incertitudes annoncées Δi , Δi_1 et Δi_2 ne sont pas des majorants de $|\varepsilon_i|$, $|\varepsilon_{i_1}|$ et $|\varepsilon_{i_2}|$ mais seulement les demi-largeurs d'intervalles ayant une probabilité $P < 1$ de contenir les valeurs exactes respectives I , I_1 et I_2 .

Dans ces conditions, la probabilité P' pour que l'intervalle $[-(\Delta i + \Delta i_1 + \Delta i_2), +(\Delta i + \Delta i_1 + \Delta i_2)]$ contienne la différence $(\hat{i}_1 + \hat{i}_2 - \hat{i})$ est beaucoup plus proche de l'unité que ne l'est P (par exemple $(P' = 0,999)$ tandis que $P = 0,95$) à cause du caractère probabiliste de la composition des erreurs. Pour avoir $(\hat{i}_1 + \hat{i}_2 - \hat{i}) = (\Delta i + \Delta i_1 + \Delta i_2)$, par exemple, il faudrait en effet que les intensités \hat{i}_1 et \hat{i}_2 soient entachées d'erreurs positives respectivement voisines de Δi_1 et Δi_2 (ce qui est déjà fort improbable), tandis que \hat{i} serait une mesure de I affectée d'une erreur négative voisine de $-\Delta i$, elle-même improbable. L'ensemble a peu de chances de se produire ce qui, mathématiquement, s'écrit $(1 - P') \ll (1 - P)$.

Cela signifie aussi que l'intervalle ayant la probabilité P de contenir la valeur de $(\hat{i}_1 + \hat{i}_2 - \hat{i})$ est plus étroit que $[-(\Delta i + \Delta i_1 + \Delta i_2), +(\Delta i + \Delta i_1 + \Delta i_2)]$. On peut montrer qu'il est de la forme $[-\delta i, +\delta i]$ avec $\delta i = \sqrt{(\Delta i)^2 + (\Delta i_1)^2 + (\Delta i_2)^2} = 2,6 \text{ mA}$.

La justification de cette expression exige des développements mathématiques absolument incompatibles avec le niveau du collège ou même celui du lycée. En revanche, au niveau des classes post-baccalauréat, une fois acquis le théorème selon lequel la variance d'une somme algébrique de variables aléatoires *indépendantes* est égale à la somme des variances de ces variables [réf. (1) page 118], cette propriété est aisément justifiable.

Au niveau du collège, il n'y a pas lieu d'évoquer le calcul de l'incertitude d'une somme algébrique à partir des incertitudes de chacun des termes de la somme.

Remarque 2 : Au collège, comme en classe de 2^{de}, où les appareils ne sont généralement pas de très grande qualité, le calcul de l'incertitude sur une mesure donnée par un appareil à l'aide de sa notice est risqué. Parfois, en effet, après un ou deux ans d'usage, les appareils de bas de gamme se dérèglent quelque peu et il arrive que la valeur exacte de la grandeur mesurée par l'un d'eux soit nettement extérieure à l'intervalle calculé d'après la notice.

Prenons un exemple réel : dans un établissement, on a placé en série 20 multimètres numériques à 2 000 points, utilisés en ampèremètres : il s'agit d'un lot d'appareils identiques, utilisés depuis plusieurs

années. Pour le calibre utilisé (200 mA), la notice commune indique que l'incertitude Δi se détermine par l'expression : $\Delta i = 1 \% VL + 1 UR$.

L'intensité mesurée I étant voisine de 96, on en déduit $\Delta i = 1,1$ mA. Si les appareils étaient neufs, entre la plus grande et la plus petite de leurs indications, la différence ne devrait guère dépasser $2 \Delta i = 2,2$ mA. Or, les différentes indications des appareils, rangées par ordre croissant, et exprimées en milliampères sont :

94,1 ; 94,8 ; 94,8 ; 94,9 ; 95,0 ; 95,0 ; 95,1 ; 95,1 ; 95,1 ; 95,1 ; 95,1 ; 95,2 ; 95,4 ; 97,7 ; 98,2 ; 98,4 ; 99,6 ; 99,7 ; 100,8 ; 101,5.

On constate que l'étendue $r = \hat{i}_{\max} - \hat{i}_{\min}$ est égale à 6,4 mA : elle est donc environ trois fois plus grande que ce que laissait prévoir la notice, par conséquent ces appareils ne sont plus conformes à ce que prédit cette dernière.

Sans être toujours aussi marqué que dans l'exemple précédent, le fait est tout de même assez fréquent. La meilleure solution consisterait à réétalonner les appareils lorsque, comme précédemment, leurs performances collectives réelles contredisent la notice. Sans cela, nous verrons qu'il est tout de même possible de réévaluer globalement leur incertitude.

Remarque 3 : Les incertitudes dues aux appareils peuvent se combiner avec d'autres. Supposons que l'on utilise un ohmmètre relativement neuf : lorsqu'il affiche $R = 46,82 \Omega$ la notice indique que la précision relative de cette mesure est telle que $\Delta R = 0,5 \% VL + 1 UR$, ce qui correspond à $\Delta_1 R = (0,23 + 0,01) \Omega = 0,24 \Omega$. En principe, donc, le constructeur garantit que, s'il affiche $46,82 \Omega$, l'ohmmètre a mesuré une résistance comprise entre $46,58$ et $47,06 \Omega$ (il est plus réaliste de dire que l'intervalle $\{46,58 \Omega, 47,06 \Omega\}$ a de fortes chances de contenir la valeur exacte de R). Supposons donc ceci réalisé pour une résistance "radio" à la température de la salle. Si, comme nous l'avons vu plus haut, la température de la résistance est susceptible de varier au cours de la séance de travail, l'incertitude $\Delta_2 R$ due à cette variation, dont nous avons vu qu'elle est voisine de $0,1 \Omega$, se combine à $\Delta_1 R = 0,24 \Omega$. Sauf cas particulier, le signe de la variation est *a priori* imprévisible, aussi conçoit-on que l'incertitude finale, ΔR , soit supérieure à la fois à $\Delta_1 R$ et à $\Delta_2 R$.

On en restera là. On ne posera pas, en particulier, $\Delta R = \Delta_1 R + \Delta_2 R$, ce qui conduit au résultat $R = 46,8 \Omega \pm 0,35 \Omega$, expression dans laquelle l'incertitude ΔR est généralement trop grande.

L'expression $\Delta R = \sqrt{\Delta_1 R^2 + \Delta_2 R^2}$ qui conduit à $\Delta R = 0,27 \Omega$ est plus réaliste, mais son utilisation au niveau du collège est exclue. Il reste que le résultat $R = 46,8 \Omega \pm 0,3 \Omega$ serait certainement le bon.

Remarque 4 : Tout comme des ampèremètres montés en série n'indiquent pas tout à fait la même intensité, plusieurs thermomètres, placés dans le même bain homogénéisé, n'indiquent pas rigoureusement la même température ; plusieurs balances, soumises à la même masse ou au même poids, ne donnent pas le même résultat (parfois, une même balance donne une indication qui varie légèrement avec l'endroit où l'on place la masse à mesurer), etc.

On peut profiter de ces observations pour évaluer l'ordre de grandeur de l'incertitude qui affecte une mesure d'élève quand il utilise, dans des conditions voisines, l'un de ces appareils pris au hasard. Il est alors possible d'indiquer aux élèves, sans justification, quelle incertitude ils peuvent prendre (il ne s'agit en aucune manière de leur imposer des calculs).

Exemple : supposons que le mesurage d'une même masse m , effectuée par neuf élèves différents utilisant neuf balances différentes ($n = 9$), mais de caractéristiques nominales identiques, ait donné les résultats ordonnés suivants, exprimés en grammes :

109,5 ; 109,7 ; 109,8 ; 110,0 ; 110,0 ; 110,0 ; 110,1 ; 110,1 ; 110,3.

Nous constatons qu'aucune de ces 9 mesures ne s'écarte sensiblement des autres, aussi ne peut-on suspecter aucune des balances d'être fautive. L'étendue r de ces mesures (différence entre la plus grande mesure et la plus petite), est égale à $0,8$ g. Dans ces conditions, on peut faire admettre le résultat suivant : *lorsque l'on utilise l'une de ces balances, prise au hasard, pour mesurer une masse d'une centaine de grammes et que l'on trouve le résultat $m = m_i$, la valeur exacte de la masse mesurée a de fortes chances*

(de l'ordre de 95 sur 100) de se trouver dans l'intervalle $(m_i - 0,7 r, m_i + 0,7 r)$.

Ici, puisque $r = 0,8$ g, nous arrondissons la quantité $0,7 r$, qui représente l'incertitude absolue sur une mesure au niveau de confiance 95 % est égale à $0,56$ g ; comme il s'agit seulement de donner des ordres de grandeur, arrondissons-la à $0,6$ g : on dira qu'il y a environ 95 chances sur cent pour que, lorsqu'un élève utilise une balance de ce type prise au hasard (parmi toutes celles qui existent et pas nécessairement dans la collection du collège) et trouve le résultat m_i , la vraie valeur de m se trouve dans l'intervalle $(m_i - 0,6$ g, $m_i + 0,6$ g). Par exemple, l'élève obtenant dans ces conditions $m_3 = 109,8$ g, peut donner comme intervalle d'incertitude $(109,2$ g, $110,4$ g) ou, mieux, écrire $m_3 = (109,8 \pm 0,6)$ g.

Donnons quelques indications sur la manière d'obtenir ce résultat : nous supposons que les neuf mesures de m sont tirées d'une population gaussienne d'écart-type σ . Dans ce cas, en moyenne, l'étendue r d'un échantillon de 9 valeurs est égale à $2,98 \sigma$ aussi estime-t-on σ par $s = r / 2,98$. L'incertitude, que l'on prend généralement égale à 2 fois l'écart-type σ , peut donc être estimée par $2 r / 2,98 = 0,67 r$, soit environ $0,7 r$.

Dans le cas de $n = 6$, nous estimons 2σ par $2 r / 2,54 \approx 0,8 r$; pour $n = 12$ par $2 r / 3,26 \approx 0,6 r$.

On peut également estimer σ à partir de l'estimateur habituel $s = \sigma_{n-1}$ (voir plus loin) : l'incertitude Δm sur une mesure, au niveau de confiance 95 %, est alors égale à $2 \sigma_{n-1}$, soit, ici, $0,48$ g. On remarque que ce résultat est proche du précédent, plus simple à mettre en œuvre, car, sans arrondi, la quantité $0,67 r$ est égale à $0,54$ g.

Remarque 5 : Si les élèves et les balances ont été supposés différents, c'est pour pouvoir considérer que les neuf mesures obtenues sont *indépendantes* : dans le cas contraire (neuf élèves mesurant successivement la masse d'un même objet avec une même balance, ou un seul élève, pouvant être très malhabile ou très adroit, effectuant neuf mesures avec des balances différentes, etc.), elles seraient plus ou moins fortement "liées", ou, en langage savant, *corrélées*. Or le coefficient $0,7$ intervenant ci-dessus suppose que les mesures sont indépendantes et, *a priori*, de même qualité. S'il n'en est pas ainsi, l'intervalle d'incertitude doit être plus ou moins élargi.

Autre exemple de mesures corrélées : une longueur L peut être considérée comme la somme de deux longueurs L_1 et L_2 . Deux élèves mesurent L_1 et obtiennent deux résultats λ'_1 et λ''_1 . Ils font alors appel à un troisième pour mesurer L_2 et celui-ci trouve λ_2 . Les sommes $\lambda' = \lambda'_1 + \lambda_2$ et $\lambda'' = \lambda''_1 + \lambda_2$ constituent bien deux mesures de $L = L_1 + L_2$, mais ce sont deux mesures corrélées et non deux mesures indépendantes.

Remarque 6 : La sensibilité d'une balance varie avec la masse qu'on mesure : notre étude ayant été menée à propos d'une masse de l'ordre de 100 grammes, l'incertitude absolue de $0,6$ g que nous avons déterminée n'est valable que pour des mesures voisines (par exemple de 80 à 120 g). Pour des masses de l'ordre de 10 grammes, par exemple, ou de 500 grammes, il faudrait recommencer l'opération et les incertitudes seraient certainement différentes.

Remarque 7 : Dans le cas précis qui vient d'être décrit, la moyenne μ des neuf mesures ($109,93$ g) constitue le meilleur résultat collectif que l'on puisse donner pour la mesure de m à l'issue de l'opération de comparaison des indications des neuf balances. La notion de moyenne étant au programme de mathématiques de la classe de 4^{ème}, on pourra, dès le collège, retenir que *la moyenne de plusieurs mesures indépendantes est le meilleur estimateur de la grandeur mesurée*.

On montre par ailleurs (voir plus loin) que la vraie valeur de la masse mesurée à l'aide des neuf balances précédentes, a de fortes chances (95 chances sur cent) d'appartenir à l'intervalle $(109,93$ g \pm $0,17$ g), soit $(109,76$ g ; $110,10$ g) : *la mise en commun de plusieurs résultats indépendants améliore la précision*.

Indiquons, pour les seuls professeurs, ce sur quoi repose ce résultat :

- estimation de l'écart-type de la mesure individuelle de m : $\sigma_{n-1} = 0,224$ g ;

- estimation de l'écart-type de la moyenne μ des neuf pesées, considérée comme une nouvelle variable aléatoire :

$$\sigma_{\mu} = \frac{\sigma_{n-1}}{\sqrt{9}} = \frac{0,224 \text{ g}}{3} = 0,075 \text{ g} ;$$

- calcul de la demi-largeur de l'intervalle de confiance de μ au niveau de confiance 95% :

$$\Delta\mu = t \frac{\sigma_{n-1}}{\sqrt{9}} = 2,31 \times 0,075 \text{ g} = 0,17 \text{ g} .$$

Dans cette expression, t est le coefficient de Student correspondant à 9 mesures et $P = 0,95$ (cf. annexe 4).

La plupart des considérations qui précèdent sont bien entendu au seul usage des professeurs. Au cours de cette première étape de sensibilisation aux résultats de mesure, il importe surtout d'être attentif à la notion de chiffres significatifs et à la cohérence des résultats de certains calculs. A l'occasion de telle ou telle manipulation, on pourra mettre en évidence que certaines grandeurs sont définies avec une certaine imprécision (hauteur d'une colonne de liquide dans un tube, distance focale d'une lentille en lumière blanche...), tandis que d'autres peuvent légèrement varier en fonction de paramètres physiques non maîtrisés : on en déduira le nombre de chiffres significatifs à conserver. Mais ceci ne fera pas l'objet d'un cours. De même c'est lorsque l'occasion se présentera, c'est-à-dire peut-être une ou deux fois dans l'année, que l'on effectuera plusieurs mesures indépendantes d'une même grandeur dans le but de mettre en évidence la dispersion naturelle des mesures. Une telle occasion permet *au professeur* d'évaluer l'incertitude sur une mesure individuelle. En outre, si la grandeur mesurée présente de l'intérêt, il faut en profiter pour l'estimer par la *moyenne* des mesures, celles-ci étant indépendantes et effectuées dans les mêmes conditions : on sera généralement surpris de la précision d'une telle évaluation.

Quelques précisions supplémentaires (ou rappels) pour les professeurs

Les erreurs les plus petites sont plus probables

On rappelle que la mesure \hat{a} d'une grandeur a de valeur exacte A comporte généralement une erreur $\varepsilon_a = \hat{a} - A$, que celle-ci soit due aux appareils, au manipulateur ou à la méthode employée ; cela se traduit par une incertitude "multifactorielle" Δa sur les mesures individuelles obtenue en classe et conduit à l'écriture $a = \hat{a} \pm \Delta a$, sachant que l'intervalle $[\hat{a} - \Delta a, \hat{a} + \Delta a]$ doit avoir une probabilité P de contenir A . Cette probabilité est en général de 95 %, mais on peut désirer qu'elle soit portée à 99 %. On comprend que l'incertitude Δa est d'autant plus grande que P est proche de l'unité. Dans ce qui suit, sauf mention contraire, nous nous placerons toujours dans le cas où $P = 95 \%$.

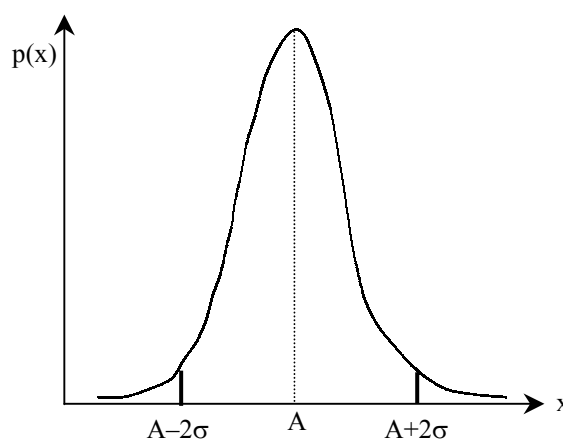


Figure 2

L'erreur ε_a est inévitable ; sa valeur absolue est, en général, nettement inférieure à Δa , puisque l'inégalité $|\varepsilon_a| > \Delta a$ n'est réalisée que dans 5 % des cas seulement (par exemple).

Le plus important, dans la distribution de Gauss, qui est celle de la répartition de la majorité des erreurs aléatoires, c'est qu'elle est décroissante de part et d'autre de sa valeur moyenne X (inconnue) qui, en l'absence d'erreur systématique, est censée représentée la valeur exacte cherchée. Autrement dit, les petites erreurs sont plus probables (et plus donc plus fréquentes) que les grandes. C'est cette notion qui est à la base de la théorie moderne des incertitudes et qui s'oppose de manière irréconciliable avec l'ancien "calcul d'incertitudes".

La notion de mesures "indépendantes"

Les théories simplifiées qui justifient nos règles et nos calculs, supposent que les différentes mesures que nous utilisons pour estimer une grandeur A et calculer son intervalle de confiance ΔA à un niveau de confiance donné P (vocabulaire voisin du précédent et synonyme), sont indépendantes. L'indépendance des mesures est une notion assez difficile à percevoir sur le terrain. De nombreux professeurs ne débusquent pas assez vite le fait que des erreurs soient liées (et donc que les mesures correspondantes ne soient pas indépendantes) et chaque fois que l'on m'a signalé que mon point de vue "ne marchait pas", c'était, en fait parce qu'au départ, les mesures n'étaient pas indépendantes.

Nous allons nous appuyer sur un exemple pour tenter d'y voir plus clair à ce sujet.

Deux tensions continues voisines V_1 et V_2 , bien définies, ont été mesurées à l'aide d'une collection de 19 multimètres de bonne qualité, ces appareils, utilisés depuis plusieurs années, n'ont jamais été recalibrés. Voici les résultats obtenus, exprimés en volts :

	V_1	V_2	$V_1 - V_2$	$V_1 + V_2$
1	16,82	15,09	1,73	31,91
2	16,91	15,17	1,74	32,08
3	16,85	15,11	1,74	31,96
4	16,97	15,23	1,74	32,20
5	16,71	14,96	1,75	31,67
6	16,84	15,11	1,73	31,95
7	16,87	15,13	1,74	32,00
8	16,95	15,21	1,74	32,16
9	16,92	15,18	1,74	32,10
10	16,89	15,15	1,74	32,04
11	16,69	14,97	1,72	31,66
12	16,74	15,03	1,71	31,77
13	16,88	15,15	1,73	32,03
14	16,86	15,12	1,74	31,98
15	16,66	14,95	1,71	31,61
16	16,83	15,10	1,73	31,93
17	16,84	15,11	1,73	31,95
18	16,87	15,13	1,74	32,00
19	17,07	15,32	1,75	32,39

Nous allons étudier ces séries de 19 mesures.

La moyenne des mesures de V_1 est 16,85 V.

Celle des mesures de V_2 est 15,12 V.

La moyenne de la différence $V_1 - V_2$ est évidemment égale à la différence des moyennes, soit 1,73 V et la moyenne de la somme est égale à la somme des moyennes, soit 31,97 V.

L'écart-type σ_{v_1} de la série des V_1 est de 97 mV ; celui de la série des V_2 , σ_{v_2} , vaut 90 mV. Mais l'écart-type σ_d de la différence $V_1 - V_2$ n'est que de 11 mV, alors que σ_s , écart-type de la somme $V_1 + V_2$ est égal à 187 mV !

Plusieurs tests (graphiques, comme ceux de Henry et de Dixon, ou numériques comme celui de Shapiro-Wilk) montrent en outre que les échantillons de 19 mesures V_1 ou V_2 peuvent être considérés comme tirés de populations normales [norme X 06-050]. C'est aussi le cas du logiciel « Incertitudes de mesure » distribué par le CNDP.

Constatons tout d'abord que les appareils étudiés peuvent encore être considérés comme bons : si l'on prend comme incertitude ΔV_1 sur une mesure unique la quantité $2 \sigma_{v_1}$, on trouve $\Delta V_1 \approx 0,2$ V, ce qui correspond à une incertitude relative $\Delta V_1 / V_1 = 1,2 \times 10^{-2}$, tandis que, dans les mêmes conditions, $\Delta V_2 / V_2 = 1,3 \times 10^{-2}$.

En revanche, l'indication de la notice, spécifiant que l'incertitude est égale à 0,15 % de la valeur affichée augmentée de la valeur correspondant à un digit, soit 35 mV pour V_1 et 33 mV pour V_2 , n'est plus d'aucune actualité : ces quantités doivent être multipliées par trois !

Nous pouvons encore observer que la différence $V_1 - V_2$ et la somme $V_1 + V_2$ sont fournies avec la même précision relative que chacune des tensions V_1 ou V_2 (l'incertitude relative sur $V_1 - V_2$, comme celle sur $V_1 + V_2$, vaut $1,3 \times 10^{-2}$). **Les mesures de V_1 et V_2 effectuées avec le même appareil, sont en effet liées** : on ne peut donc pas leur appliquer les règles de composition des incertitudes absolues des variables aléatoires indépendantes.

Au demeurant, la mesure de $V_1 - V_2$ est bien plus précise en utilisant le même appareil pour mesurer V_1 et V_2 qu'en prenant un appareil pour mesurer V_1 et un autre pour mesurer V_2 . Dans les mêmes conditions, la mesure de $V_1 + V_2$, elle, est moins précise.

Montrons-le simplement : tirons au sort, pour chaque mesure de V_1 par un appareil, le numéro de l'appareil qui donnera V_2 . Un tirage a donné le résultat suivant :

V_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
V_2	15	1	12	3	9	17	6	11	19	14	13	10	7	18	8	2	5	16	4

Ainsi la différence $V_1 - V_2$ et la somme $V_1 + V_2$ seront évaluées en prenant pour V_1 l'indication de l'appareil n°1 et pour V_2 celle de l'appareil n°15, puis on recommence ensuite avec les couples d'appareils 2 et 1, 3 et 12, etc. Autrement dit remplaçons les mesures liées V_1 et V_2 par des mesures indépendantes.

Les résultats que l'on obtient sont portés sur le tableau suivant :

	V_1	V_2	$V_1 - V_2$	$V_1 + V_2$
1- 15	16,82	14,95	1,87	31,77
2- 1	16,91	15,09	1,82	32,00
3- 12	16,85	15,03	1,82	31,88
4-3	16,97	15,11	1,86	32,08
5-9	16,71	15,18	1,53	31,89
6-17	16,84	15,11	1,73	31,95
7-6	16,87	15,11	1,76	31,98
8-11	16,95	14,97	1,98	31,92
9-19	16,92	15,32	1,60	32,24
10-14	16,89	15,12	1,77	32,01
11-13	16,69	15,15	1,54	31,84
12-10	16,74	15,15	1,59	31,89
13-7	16,88	15,13	1,75	32,01
14-18	16,86	15,13	1,73	31,99
15-8	16,66	15,21	1,45	31,87
16-2	16,83	15,17	1,66	32,00
17-5	16,84	14,96	1,88	31,80
18-16	16,87	15,10	1,77	31,97
19-4	17,07	15,23	1,84	32,30

Les écarts-types σ_{v_1} et σ_{v_2} sont toujours égaux, respectivement à 97,2 et 90,3 mV ; mais cette fois-ci, l'écart-type σ_d de la différence, qui était de 11 mV lorsque les mesures de V_1 et V_2 étaient liées, devient égal à 136 mV. En revanche, celui de la somme, σ_s , régresse de 187 mV à 130 mV.

$$\text{Or : } \sqrt{97,2^2 + 90,3^2} = 133 \text{ mV} ; \text{ on a donc maintenant : } \sigma_d \approx \sigma_s \approx \sqrt{\sigma_{v_1}^2 + \sigma_{v_2}^2} .$$

On constate donc, sur ce cas particulier, qu'il faut être très attentif au fait que des mesures soient indépendantes ou non, puisqu'en effectuant des différences de mesures indépendantes, on a multiplié par 12 l'écart-type de la différence $V_1 - V_2$ et fortement réduit celui de la somme $V_1 + V_2$. À l'évidence, les mesures V_1 et V_2 de chaque multimètre sont affectées d'erreurs voisines : l'opération différence élimine en grande partie ces erreurs ; l'opération somme, elle, les ajoute.

On constate aussi que les écarts-types σ_d et σ_s sur les séries de différences ou de sommes de deux mesures indépendantes sont très proches de la somme quadratique des écarts-types σ_{v_1} et σ_{v_2} . Dépendant du hasard, car il existe $19! \approx 1,2 \times 10^{17}$ manières d'apparier les appareils, σ_d , par exemple, peut certes varier entre les valeurs extrêmes 11 mV (lorsque V_1 et V_2 sont liées, ce qui correspond à une probabilité de 10^{-17}) et 184 mV (lorsqu'à la plus grande valeur de V_1 on fait correspondre la plus petite valeur de V_2 et vice versa, etc., situation de probabilité là encore voisine de 10^{-17}), mais, pour la plupart des combinaisons possibles, sa valeur est proche de 0,13 V.

Un autre appariement, dû lui aussi au hasard, conduisant aux couples d'appareils (1,4) ; (2,2) ; (3,17) ; (4,15) ; (5,1) ; (6,9) ; (7,6) ; (8,16) ; (9,5) ; (10,18) ; (11,8) ; (12,12) ; (13,7) ; (14,10) ; (15,11) ; (16,14) ; (17,13) ; (18,19) et

(19,3) conduit à $\sigma_d = 134$ mV et $\sigma_s = 131$ mV, ce qui est encore plus proche de la relation théorique qui correspond aux grands nombres de mesures.

C'est l'addition quadratique des écarts-types, jointe à l'hypothèse de normalité, qui conduit à l'addition quadratique des incertitudes absolues dans le cas d'une somme ou d'une différence de grandeurs indépendantes.

Intérêt de la notion d'écart-type

La notion d'écart-type permet de comparer, par exemple, les indications de deux séries de multimètres : supposons que le professeur dispose d'une ligne sèche permettant d'alimenter chaque poste de travail avec un même générateur de tension. Chaque groupe d'élèves (repéré par le nombre n variant par exemple de 1 à 9) dispose de deux multimètres, un appareil A_n "grand public", de faible coût, et un véritable appareil de mesure, B_n .

On relève les 9 indications A_n , les 9 indications B_n et on calcule les écarts-types σ_A et σ_B .

Si le premier de ces écarts-types est comparable au second (si σ_A ne dépasse pas l'ordre de grandeur $1,5 \sigma_B$), on se réjouit d'avoir des appareils peu coûteux et pourtant de bonne qualité. Mais si $\sigma_A > 2,5 \sigma_B$, on peut sans grand risque de se tromper, affirmer que les appareils de type A sont moins précis que ceux du type B (des tables, dites de Fisher, permettent de décider, en fonction de critères quantitatifs, si le rapport des écarts-types peut être considéré comme significatif, nous en dirons un mot en annexe 8).

Il peut arriver que l'on soit en mesure de mesurer collectivement une même grandeur à l'aide d'appareils différents et par des élèves différents (on obtient ainsi un nombre n de mesures indépendantes voisin de 10) : on calcule alors la moyenne m des n mesures et leur écart-type σ . Ces résultats permettent de répondre à plusieurs questions :

a) La moyenne m est alors le meilleur estimateur de la valeur de la grandeur mesurée.

b) On peut présenter l'intervalle $(m - 2 \sigma, m + 2 \sigma)$ comme l'intervalle de confiance d'une mesure individuelle, à un niveau voisin de 95 %.

Nous pouvons en effet reprendre à notre compte la phrase un peu vague qui était insérée dans les programmes précédents de mathématiques, selon laquelle, "dans de nombreux phénomènes, le pourcentage de mesures appartenant à l'intervalle $(m - 2 \sigma, m + 2 \sigma)$ ou $(m - 3 \sigma, m + 3 \sigma)$ est voisin de 95 % ou 99 %".

Supposons que nous disposions de nombres au hasard tirés d'une population gaussienne centrée sur la valeur X et d'écart-type σ . Constituons des échantillons de 9 nombres (par exemple) pris au hasard : $x_1, x_2, x_3, \dots, x_9$; calculons pour chacun de ces échantillons l'estimateur σ_{n-1} qui sera défini plus loin, puis cherchons si les intervalles $[x_i - 1,96 \sigma_{n-1} ; x_i + 1,96 \sigma_{n-1}]$ contiennent X . En général, la réponse est positive, mais il arrive que pour une valeur x_i , parfois deux, ce ne soit pas le cas. Re commençons l'opération un grand nombre de fois, par exemple 100 fois. On constate alors que sur 900 nombres utilisés ($9 \times 100 = 900$), 45 intervalles environ, soit 5 % d'entre eux, ne contiennent pas X .

C'est ce résultat, avéré, qui a été transposé ci-dessus au cas des erreurs de mesures physiques, en remplaçant le coefficient 1,96 par 2 et σ_{n-1} par " σ ". Rappelons encore que cette transposition suppose que les erreurs se répartissent selon une loi gaussienne et que le procédé de mesure ne comporte pas d'erreur systématique.

c) On peut aussi se demander quel est l'ordre de grandeur de l'incertitude qui affecte, non pas une mesure unique, comme précédemment, mais la moyenne m de n mesures indépendantes. L'idée importante, c'est que si les mesures individuelles constituent des variables aléatoires de moyenne A et d'écart-type σ_a , alors la moyenne de n mesures indépendantes est elle-même une variable aléatoire de même moyenne A (on dit plutôt d'espérance mathématique A) et d'écart-type $\sigma_m = \frac{\sigma_a}{\sqrt{n}}$.

Dans ces conditions, si n est supérieur à 20, le fait de prendre comme estimateur d'une grandeur mesurée, non pas une mesure unique mais la moyenne m des n mesures indépendantes de même type, améliore la précision d'un facteur proche de \sqrt{n} . Dans le cas où n est nettement inférieur à 20, ce facteur est un peu inférieur à \sqrt{n} (voir l'annexe 4) mais on peut retenir que :

L'incertitude sur la moyenne m , à un niveau de confiance de 95%, est proche de $\frac{2\sigma}{\sqrt{n}}$

d) Plaçons une dizaine d'ampèremètres en série (ou des voltmètres en parallèle) et relevons les indications de ces appareils. On obtient une série statistique dont on calcule l'écart-type σ_n . Il est utile de comparer σ_n à l'incertitude annoncée par le constructeur : souvent celle-ci est voisine de $2 \sigma_n$, ce qui correspond à un niveau de confiance P proche de 95 %.

Mais, comme nous l'avons vu précédemment, cette incertitude annoncée est parfois nettement insuffisante. Il arrive qu'elle soit inférieure à σ_n ($P < 68 \%$). Pour des appareils de qualité, en revanche, elle représente parfois plus de $3 \sigma_n$ ($P > 99,7 \%$) : dans ce cas elle est pratiquement un majorant de la valeur absolue de l'erreur (ce qui ne rend pas pour autant légitime les calculs d'incertitudes d'autrefois).

Ce qui peut être présenté en première et terminale scientifiques ou industrielles

Réflexions et activités spécifiques à la classe de 1^{ère}

Il faut expliquer fermement aux élèves que la possibilité de répéter des mesures ne doit en aucune manière constituer un prétexte à de mauvaises mesures initiales ou à l'utilisation de méthodes peu précises, car cette répétition, coûteuse et consommatrice de temps, ne permet souvent pas de rattraper la précision perdue au départ.

En utilisant leur calculatrice, les élèves peuvent aussi comprendre, dans un cas particulier, comment les erreurs entachant les mesures primaires se propagent dans les calculs (il ne peut s'agir que d'une sensibilisation, si le niveau de la classe le permet).

Donnons un exemple : en chauffant un mélange de cuivre de masse m_1 et de poudre de soufre (en excès), les élèves obtiennent un produit d'aspect métallique de masse m_2 dont ils doivent déterminer la formule stœchiométrique Cu_nS (en supposant qu'elle existe).

Supposons qu'un élève prenne $m_1 = 2,00$ g de cuivre. Si le sulfure formé est bien Cu_2S , il devrait trouver $m_2 = 2,505$ g. Or le coefficient stœchiométrique n est donné par l'expression :

$$n = \frac{32,1}{63,5} \times \frac{m_1}{m_2 - m_1}$$

Le professeur, lui, sait, à partir de cette expression, établir la relation :

$$\frac{dn}{n} = \frac{m_2}{m_2 - m_1} \times \left(\frac{dm_1}{m_1} - \frac{dm_2}{m_2} \right) \approx 5 \times \left(\frac{dm_1}{m_1} - \frac{dm_2}{m_2} \right)$$

Mais, par de petits calculs de variations effectués à la machine, en remplaçant m_1 et m_2 par des valeurs proches, un élève de première, de son côté, peut très bien *constater* que des erreurs δm_1 et δm_2 commises sur m_1 et m_2 (respectivement), entraînent une erreur δn sur n telle que :

$$\frac{\delta n}{n} \approx 5 \times \left(\frac{\delta m_1}{m_1} - \frac{\delta m_2}{m_2} \right)$$

Par exemple, si $m_1 = 2,02$ g au lieu de 2,00 g, on a $\delta m_1 / m_1 = 1 \%$. En remplaçant m_1 par cette valeur dans l'expression donnant n (et en conservant $m_2 = 2,505$ g), on trouve $n = 2,10$ ce qui correspond à une variation δn telle que $\delta n / n = 5 \%$ = $5 \delta m_1 / m_1$. De même, pour $m_1 = 2,00$ g et $m_2 = 2,530$ g

($\delta m_2 / m_2 = 1 \%$), on trouve $n = 1,907$, soit une variation δn telle que $\delta n / n \approx -5 \% = -5 \delta m_2 / m_2$.

L'expression précédente montre donc qu'il faut être soigneux dans les mesures de m_1 et m_2 car une erreur de 2,5 % par excès sur m_1 combinée à une erreur par défaut de 2,5 % sur m_2 , par exemple, entraîne une valeur de n supérieure à 2,5 et donc une erreur sur le coefficient stœchiométrique n , car on prendra alors $n = 3$ au lieu de $n = 2$.

Heureusement, les mesures de m_1 et de m_2 étant effectuées avec la même balance, nous avons vu, à propos des voltmètres (§ 3.2), qu'elles sont corrélées par l'appareil : $\frac{\delta m_1}{m_1}$ et $\frac{\delta m_2}{m_2}$, assez fortement liées, ont ainsi de fortes chances d'être de même signe.

Cela explique le succès de cette manipulation qui est souvent réalisée en classe.

La simulation numérique, à l'aide de nombres au hasard à répartition gaussienne, peut être utilisée pour familiariser les élèves avec certains phénomènes.

Supposons par exemple que l'on dispose de mille nombres x_i répartis de manière normale autour de zéro, avec un écart-type σ égal à l'unité. Pour obtenir des nombres répartis autour de $X = 85$ avec $\sigma' = 3$, par exemple, il suffit de transformer les nombres x_i en y_i tels que $y_i = 85 + 3 x_i$: on peut donc obtenir n'importe quelle population gaussienne.

On peut ainsi simuler un produit de grandeurs U et I affectées d'erreurs aléatoires à répartition gaussienne comme $P = U I$, ou un quotient comme $R = \frac{U}{I}$ et comparer $\frac{\sigma_P}{P}$ ou $\frac{\sigma_R}{R}$, quantités calculées *a posteriori* sur une centaine de cas, à $\frac{\sigma_U}{U}$ et $\frac{\sigma_I}{I}$ quantités connues.

$$\text{On constatera ainsi que : } \frac{\sigma_P}{P} = \frac{\sigma_R}{R} = \sqrt{\left(\frac{\sigma_U}{U}\right)^2 + \left(\frac{\sigma_I}{I}\right)^2}.$$

Ce résultat particulier étant indépendant du type de distribution utilisée, il n'est pas nécessaire que les nombres au hasard considérés soient tirés d'une distribution gaussienne ; des nombres issus d'une répartition uniforme sur un intervalle donné conviennent aussi bien.

Ce qui devrait être acquis en terminale

Variable aléatoire, estimation de sa moyenne

La notion de variable aléatoire figure au programme de mathématiques, avec ses corollaires : fonction de répartition, écart-type, etc. On en profite pour présenter, en utilisant un vocabulaire correct, le caractère probabiliste de certaines erreurs : si l'on mesure une grandeur X , inconnue, à l'aide d'une méthode ne comportant pas d'erreur systématique, chaque mesure est assimilable à la valeur prise par une variable aléatoire x centrée sur X , c'est-à-dire que sa moyenne, ou espérance mathématique $E(x)$, vérifie l'égalité $E(x) = X$.

Un ensemble, appelé échantillon de n mesures indépendantes x_i , obtenues en utilisant la même méthode et du matériel comparable, une fois débarrassé des mesures manifestement fausses (TP "collectif"), permet d'estimer X par la moyenne m des x_i : c'est le meilleur estimateur de X (voir annexe 2).

Par exemple, la détermination de la masse de Jupiter par l'analyse du mouvement de ses satellites, à partir de mesures de distances effectuées sur des photocopies de documents annotés (afin de déterminer le diamètre de chaque orbite de satellite et sa période de révolution) conduit à des résultats individuels qui peuvent être décevants pour tel ou tel élève. En revanche, pour un groupe d'une douzaine d'élèves travaillant indépendamment, le résultat collectif devient satisfaisant car l'erreur relative de la détermination collective peut être inférieure à 5 %.

Estimation de l'écart-type

On peut dire aux élèves que la théorie mathématique permet de montrer que le meilleur *estimateur* de l'écart-type de la variable aléatoire x n'est pas σ , l'écart-type de la série statistique des x_i obtenue (parfois appelé σ_n), mais σ_{n-1} (ou s_n) qui figure également dans la liste des fonctions pré-programmées des calculatrices de poche, d'où l'emploi de cette grandeur (*se reporter aux annexes 2 et 3*).

Intervalle de confiance

Lorsque la séance de travaux pratiques se prête à une exploitation collective des résultats et qu'une même grandeur a été mesurée par plusieurs groupes d'élèves (n), le calcul de la moyenne m des mesures et de l'estimateur σ_{n-1} de leur écart-type permet d'aller plus loin dans le seul but de conclure la séance par une estimation vraisemblable de l'incertitude de la mesure collective.

Cela conduit souvent à proposer, pour cette moyenne, un intervalle de confiance qui est généralement assez réduit pour que les élèves considèrent leur travail collectif comme gratifiant.

Dans l'hypothèse où toute erreur systématique a été écartée et où les mesures individuelles x_i sont réparties selon une loi gaussienne, on établit en effet en mathématiques que, pour un niveau de confiance donné (95 % ou 99 %), l'intervalle de confiance de la grandeur X mesurée en commun est de la forme (*annexe 4*) :

$$\left[m - t_n \frac{\sigma_{n-1}}{\sqrt{n}}, m + t_n \frac{\sigma_{n-1}}{\sqrt{n}} \right]$$

Le coefficient t_n est appelé coefficient de Student ; il dépend du niveau de confiance choisi et du nombre n ; ainsi, pour n variant de 7 à 12, t_n est un peu plus grand que 2 pour un niveau de confiance de 95 % et un peu plus grand que 3 pour un niveau de confiance de 99 %, comme le montre le tableau ci-dessous où figurent, pour des valeurs du nombre n des mesures indépendantes, les valeurs de t correspondant à deux niveaux de confiance :

n	3	4	5	6	7	8	9	10	12	14	16	20
$t_{95\%}$	4,30	3,18	2,78	2,57	2,45	2,37	2,31	2,26	2,20	2,16	2,13	2,09
$t_{99\%}$	9,93	5,84	4,60	4,03	3,71	3,50	3,36	3,25	3,11	3,01	2,95	2,86

Remarque : il est exclu que cette estimation donne l'occasion d'un exposé mathématique et encore moins qu'elle fasse l'objet d'une mémorisation. Ce qui importe, c'est de remarquer ce qui, après tout, relève du bon sens, que pour un même nombre n de mesures indépendantes, le coefficient t_n augmente avec le niveau de confiance souhaité.

Quant à la dépendance de la largeur de l'intervalle de confiance par rapport à n , elle vient essentiellement non du coefficient de Student, mais comme nous l'avons déjà vu, du facteur \sqrt{n} qui divise σ_{n-1} .

Sachant, en effet que l'écart-type σ_m de notre résultat de mesure, m , est estimé par la quantité $\frac{\sigma_{n-1}}{\sqrt{n}}$, on constate que pour $n > 10$, l'intervalle de confiance au niveau de confiance 95 % défini ci-dessus, pour lequel t_n est alors très voisin de 2, est bien celui que prévoit, pour les incertitudes de mesures industrielles, la norme ISO 9000, à savoir $m \pm 2 \sigma_m$.

Mais, justement, dans la norme ISO 9000, on recommande de calculer σ_{n-1} " à partir d'un minimum de 10 mesures de « répétabilité » de l'ensemble du processus de mesure, dans le cas de variables indépendantes ".

C'est donc uniquement parce que nous, en classe, ne pouvons pas toujours disposer d'un minimum de 10 mesures que nous faisons appel aux coefficients de Student. Il est tout à fait possible de donner les valeurs de ces coefficients en début d'année scolaire, en faisant remarquer aux élèves que l'on pourrait s'en passer si l'on disposait toujours d'une vingtaine de mesures indépendantes.

En terminale, l'usage du logiciel « Incertitudes de mesure » édité par le CNDP, qui reprend exactement les idées exposées ci-dessus, simplifie beaucoup leur mise en œuvre.

Les chiffres significatifs à conserver

On pourrait aider les élèves à exprimer un résultat final sous la forme $a = \hat{a} \pm \Delta a$, en leur proposant les critères suivants (annexe 5) :

- la suppression d'un chiffre significatif sur Δa et l'arrondi correspondant ne doit pas entraîner pour cette incertitude une modification supérieure à 4 % (ou 5 % à la rigueur).
- la suppression d'un chiffre significatif sur \hat{a} et l'arrondi correspondant ne doit pas entraîner une translation de l'intervalle de confiance supérieure à $0,2 \Delta a$.

D'après la première de ces règles, l'incertitude sur une mesure s'exprime toujours avec au plus deux chiffres significatifs ; la deuxième qui, comme la première, se justifie par des considérations de niveau de confiance, conforte le bon sens : si le fait de supprimer un chiffre significatif entraîne une erreur d'écriture égale au cinquième de l'incertitude, c'est payer trop cher la simplification recherchée.

Mais comme cela peut paraître encore un peu compliqué on peut simplifier ces critères sans trop les dénaturer et adopter le comportement suivant :

L'incertitude Δa s'exprime avec deux chiffres significatifs (au maximum) ; les derniers chiffres significatifs conservés pour l'estimateur \hat{a} sont ceux sur lesquels porte Δa .

Ainsi le résultat $m = (98,486 \pm 1,573)$ g s'écrira : $m = (98,5 \pm 1,6)$ g.

Annexes

Annexe 1 : variables aléatoires indépendantes

Soit une variable aléatoire x_1 , qui peut être le résultat de la mesure d'une grandeur X parfaitement définie par une certaine méthode M_1 . Si cette méthode ne comporte pas d'erreur systématique, $E(x_1) = X$ où $E(x_1)$ est l'espérance mathématique de x_1 que l'on peut approcher en faisant la moyenne d'un très grand nombre N de tirages successifs de x_1 .

Posons $\varepsilon_1 = x_1 - X$: ε_1 est une variable aléatoire centrée, d'espérance mathématique nulle : $E(\varepsilon_1) = 0$.

Procédons de même pour une deuxième variable aléatoire x_2 telle que $E(x_2) = X$, x_2 peut être la mesure de X par une deuxième méthode M_2 . Là encore, posons $\varepsilon_2 = x_2 - X$:

Les variables aléatoires x_1 et x_2 sont indépendantes si $E(\varepsilon_1 \times \varepsilon_2) = 0$.

En pratique, si les expressions de x_1 et de x_2 ne comportent pas de terme commun, il s'agit de variables aléatoires indépendantes (conduisant donc à des mesures indépendantes). En revanche, si on peut écrire, par exemple, $x_1 = a + b$ et $x_2 = a + c$ (cf. remarque 5, § 2.), les termes a , b et c étant eux-mêmes des variables aléatoires, alors x_1 et x_2 sont (partiellement) corrélées.

De même, soit, avec des notations évidentes, $R_1 = \frac{U_1}{I_1}$, $R_2 = \frac{U_1}{I_2}$, $R_3 = \frac{U_2}{I_1}$ et $R_4 = \frac{U_2}{I_2}$.

Seules deux de ces mesures sont indépendantes (par exemple R_1 et R_4) ; les deux autres leur sont partiellement corrélées.

Remarque : Si x_1 et x_2 sont indépendantes et d'écart-types respectifs σ_1 et σ_2 , l'écart-type σ_d de la différence $x_d = x_1 - x_2$ et celui de la somme $x_s = x_1 + x_2$ sont tous deux égaux à $\sqrt{\sigma_1^2 + \sigma_2^2}$. C'est un excellent test pour débusquer des corrélations entre variables.

Annexe 2 : les estimateurs

Un estimateur comme la moyenne m d'un échantillon ou l'écart-type σ_n de cet échantillon est lui-même une variable aléatoire, puisque, si l'on change d'échantillon, on trouve en général une autre valeur pour chacune de ces grandeurs.

On demande avant tout à un estimateur d'être "sans biais" : ceci est réalisé lorsque son espérance mathématique (ou moyenne calculée à partir d'un grand nombre d'échantillons) est précisément égale à la grandeur estimée. Ce n'est pas le cas de l'estimateur σ_n dont nous verrons dans l'annexe 3 que la moyenne sur un grand nombre d'échantillons ne donne pas σ , l'écart-type de la variable aléatoire initiale, mais une valeur plus petite : σ_n est donc un estimateur biaisé.

D'où l'emploi de σ_{n-1} qui, lui, est considéré comme sans biais (en toute rigueur, c'est son carré qui est sans biais). Il s'obtient en divisant, avant d'en prendre la racine carrée, la somme des carrés des écarts à la moyenne m , non par n , comme pour σ_n , mais par $n-1$.

La moyenne arithmétique m d'une série de n mesures x_i tirées d'une population considérée comme infinie et de moyenne générale X , est un estimateur sans biais de X : elle a même espérance mathématique X que les x_i .

On dit que m est le meilleur estimateur de X car l'autre qualité demandée à un estimateur est d'avoir un écart-type aussi faible que possible. Or, de tous les estimateurs sans biais de X que l'on peut réaliser à l'aide de combinaisons linéaires et homogènes de n variables primaires x_i identiques, la moyenne arithmétique m est celui dont l'écart-type est le plus faible ($\sigma_m = \frac{\sigma}{\sqrt{n}}$).

Annexe 3 : le cas particulier de σ_{n-1}

Il n'y a rien de magique dans le fait de remplacer n par $n-1$ au dénominateur de l'expression qui donne l'expression de l'écart-type d'un échantillon en vue d'obtenir un estimateur sans biais de l'écart-type σ . Rien non plus d'absolument parfait, comme nous allons le voir.

Soit x_i les valeurs prises par n variables aléatoires gaussiennes indépendantes d'espérance mathématique X et d'écart-type σ , lors d'un tirage unique de chacune d'elles. Nous désignons par m leur moyenne : celle-ci, nous le savons, est elle-même une variable aléatoire d'espérance mathématique X et d'écart-type $\frac{\sigma}{\sqrt{n}}$.

Nous nous intéressons à la quantité $S^2 = \sum (x_i - m)^2$.

$$S^2 = \sum x_i^2 - 2m \sum x_i + n m^2 ; \text{ or } \sum x_i = n m \text{ par définition de } m ; \text{ donc } S^2 = \sum x_i^2 - n m^2 .$$

Mais $x_i = X + \varepsilon_i$, avec $E(\varepsilon_i^2) = \sigma^2$ et $m = X + \varepsilon_m$, avec $E(\varepsilon_m^2) = \sigma^2/n$.

$$\text{Donc : } S^2 = n X^2 + 2 X \sum \varepsilon_i + \sum \varepsilon_i^2 - n X^2 - 2 n X \varepsilon_m - n \varepsilon_m^2 .$$

$$\text{Soit encore : } S^2 = 2 X \sum \varepsilon_i + \sum \varepsilon_i^2 - 2 n X \varepsilon_m - n \varepsilon_m^2 .$$

Calculons l'espérance mathématique de S^2 , notée $E(S^2)$, autrement dit, la moyenne de cette quantité quand on recommence l'opération d'échantillonnage un grand nombre de fois :

$$E(S^2) = 2 X \sum E(\varepsilon_i) + \sum E(\varepsilon_i^2) - 2 n X E(\varepsilon_m) - n E(\varepsilon_m^2) = \sum E(\varepsilon_i^2) - n E(\varepsilon_m^2) .$$

En effet $E(\varepsilon_i) = E(\varepsilon_m) = 0$ puisque $E(x_i) = E(m) = X$.

En remplaçant $E(\varepsilon_i^2)$ et $E(\varepsilon_m^2)$ par leurs valeurs respectives σ^2 et σ^2/n on trouve donc :

$$E(S^2) = n \sigma^2 - \sigma^2 = (n - 1) \sigma^2 .$$

La quantité qui a pour espérance mathématique σ^2 n'est donc pas $\frac{S^2}{n}$ mais $\frac{S^2}{n-1}$.

En toute rigueur, c'est donc la variance σ^2 dont l'estimateur sans biais est $\frac{S^2}{n-1} = \frac{\sum_{i=1}^n (x_i - m)^2}{n-1}$.

Mais $\sigma_{n-1} = \sqrt{\frac{\sum_{i=1}^n (x_i - m)^2}{n-1}}$, à défaut d'être vraiment sans biais, est le meilleur estimateur de σ .

Remarque : cette démonstration est indépendante de la loi de répartition des variables aléatoires x_i .

Annexe 4 : le coefficient t de Student

Rappelons que si les x_i sont n variables aléatoires gaussiennes de même espérance mathématique X et de même écart-type σ , leur moyenne arithmétique m est elle-même une variable aléatoire gaussienne, d'espérance mathématique X et d'écart-type $\frac{\sigma}{\sqrt{n}}$.

Quant au coefficient $t = \frac{X - m}{\frac{\sigma_{n-1}}{\sqrt{n}}}$, il suit une loi de Student à $n-1$ degrés de liberté.

Ce sont ces résultats mathématiques qui fondent la règle de calcul d'un l'intervalle de confiance centré sur la moyenne m pour un niveau de confiance donné, car la définition de t équivaut à $X = m \pm t \frac{\sigma_{n-1}}{\sqrt{n}}$.

On trouvera un tableau des valeurs de la variable aléatoire t de Student par exemple dans le recueil de normes de la statistique de l'AFNOR (norme NF 06 041). La loi de Student est en effet essentielle pour les échantillons d'effectifs inférieurs ou égaux à 16. Elle fut découverte par Gosset, lequel publia ses travaux sous le pseudonyme de "Student" au début du XX^e siècle.

Nous redonnons un tableau simplifié des t correspondant aux niveaux de confiance 95 % et 99 %, c'est-à-dire aux probabilités 0,95 et 0,99 pour l'intervalle, de contenir la vraie valeur cherchée X :

n	3	4	5	6	7	8	9	10	12	14	16	20
t _{95%}	4,30	3,18	2,78	2,57	2,45	2,37	2,31	2,26	2,20	2,16	2,13	2,09
t _{99%}	9,93	5,84	4,60	4,03	3,71	3,50	3,36	3,25	3,11	3,01	2,95	2,86

Remarque : on entend parfois dire la sottise suivante : les grands échantillons suivent la loi de Gauss, tandis que les petits obéissent à la loi de Student, ou bien celle-ci : les résultats individuels de la mesure d'une grandeur suivent une loi de Gauss mais la moyenne d'un petit nombre de résultats obéit à une loi de Student. La première de ces assertions ne veut rien dire : quelle est la variable aléatoire considérée ? La seconde est erronée.

Redisons encore que si une variable aléatoire x suit une loi de Gauss d'espérance mathématique X et d'écart-type σ , la moyenne m de n tirages indépendants de x suit aussi une loi de Gauss de même espérance mathématique X et d'écart-type réduit $\frac{\sigma}{\sqrt{n}}$ et ceci, quel que soit n .

En revanche la variable $t = \frac{X - m}{\frac{\sigma_{n-1}}{\sqrt{n}}}$ suit toujours une loi de Student à $n-1$ degrés de liberté mais, lorsque

n atteint ou dépasse 20, par exemple, une telle loi est très proche d'une loi de Gauss réduite (d'écart-type 1) et centrée (d'espérance mathématique nulle).

Annexe 5 : les chiffres significatifs

Supposons que l'on ait mesuré une force et trouvé dans un premier temps : $F = (2,56712 \pm 0,01283) \text{ N}$.

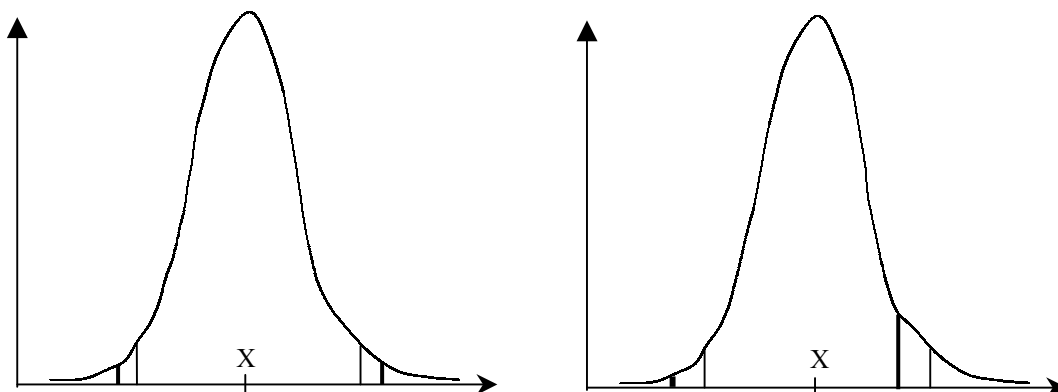
On se place dans le cas où l'estimateur de F suit une loi gaussienne, mais les résultats auxquels nous aboutirons sont tout à fait transposables à la loi de Student, très proche de la loi de Gauss pour $n \approx 9$.

L'incertitude $\Delta F = 0,01283 \text{ N}$ est donnée au niveau de confiance 95 % donc : $\sigma_F \approx 6,4 \times 10^{-3} \text{ N}$.

Comment donner ce résultat sous une meilleure forme ?

Peut-on écrire par exemple, comme cela est parfois préconisé, sous prétexte que le troisième chiffre après la virgule n'est pas certain : $F = (2,6 \pm 0,01) \text{ N}$?

Pour répondre, il faut observer les changements que provoquent la suppression des chiffres significatifs supposés en excès et les arrondis correspondants :



- Enlever le dernier 3 de ΔF , et écrire $\Delta F = 0,0128 \text{ N}$, c'est diminuer ΔF de $2,3 \times 10^{-3}$ en valeur relative. Poursuivre, et écrire $\Delta F = 0,013 \text{ N}$, c'est augmenter ΔF de $1,6 \times 10^{-2}$ en valeur relative. Donner $\Delta F = 0,01 \text{ N}$ c'est, cette fois, diminuer ΔF de 23×10^{-2} en valeur relative.

Que peut-on se permettre ? Observons tout d'abord que le standard européen recommande de prendre $\Delta F = 2 \sigma_F$ alors que les tables de la loi normale de Gauss indiquent que l'intervalle de confiance à 95 % de cette variable a pour demi-largeur $1,96 \sigma_F$; ceci revient à majorer σ_F de 2 % en valeur relative et à confondre $P = 95 \%$ et $P' = 95,4 \%$.

- D'une manière générale, si l'on minore $\sigma_{\hat{a}}$ de 5 %, sans décentrer l'intervalle, on trouve (il faut pour cela prendre une table de la fonction de répartition de Gauss), que la probabilité P passe de 95,0 % à 93,7 % : nous nous fixerons cette limite. *La suppression d'un chiffre significatif sur Δa et l'arrondi correspondant ne devra pas entraîner une modification de $\sigma_{\hat{a}}$ supérieure à 5 %*. Pour notre exemple, cela implique d'écrire $\Delta F = 0,013 \text{ N}$.

La règle précédente est assez proche de celle-ci : "on exprimera ΔF avec deux chiffres significatifs". On constatera en effet que passer de trois à deux chiffres significatifs ne provoque jamais une variation du nombre considéré de plus de $\pm 5 \%$.

- De même, écrire $F = 2,5671 \text{ N}$, c'est décentrer l'intervalle de $3,1 \times 10^{-3} \sigma_F$; supprimer le 1 provoque un nouveau décentrage de $1,6 \times 10^{-2} \sigma_F$. Enfin, écrire $F = 2,57 \text{ N}$ c'est provoquer un nouveau décentrage de $0,47 \sigma_F$.

- D'une manière générale, un décentrage de $0,4 \sigma_{\hat{a}}$ fait passer la probabilité P de 95,0 % à 93,8 % ce qui semble une limite à ne pas franchir. Si l'on adopte ce critère, on utilisera la règle suivante :

La suppression d'un chiffre significatif sur \hat{a} et l'arrondi correspondant ne devra pas entraîner une translation de l'intervalle de confiance supérieure à $0,4 \sigma_{\hat{a}}$.

Pour notre exemple, le passage de 2,567 N à 2,57 N entraîne un trop fort décentrage.

Avec nos critères, le bon résultat s'énonce donc : $F = (2,567 \pm 0,013) \text{ N}$.

On remarque que cela correspond à conserver pour l'estimation de F tous les chiffres significatifs qui sont affectés par ΔF et rien que ceux-ci. C'est une situation assez générale qui peut, sans grand danger, être érigée en règle simple :

On conserve pour F les chiffres significatifs qui interviennent dans ΔF .

Annexe 6 : incompatibilité des points de vue "classique" et probabiliste

Si $z = x + y$, et si les erreurs ε_x et ε_y commises respectivement sur les mesures x et y sont indépendantes, le point de vue classique, où Δx , Δy et Δz sont les majorants respectifs de $|\Delta x|$, $|\Delta y|$ et $|\Delta z|$, entraîne l'égalité $\Delta z = \Delta x + \Delta y$.

Du point de vue probabiliste, supposons que Δx et Δy soient respectivement les incertitudes au niveau de confiance 95 % des mesures de x et de y : $\Delta x = 2 \sigma_x$; $\Delta y = 2 \sigma_y$.

La quantité $\Delta x + \Delta y$ représente donc $2 (\sigma_x + \sigma_y)$.

Or nous avons vu que si x et y sont indépendantes, $\sigma_z = \sqrt{\sigma_x^2 + \sigma_y^2}$.

Plaçons-nous dans le cas d'école où $\sigma_x = \sigma_y = \sigma$. On a alors $\sigma_z = \sigma \sqrt{2}$.

Dans ce cas, $\Delta x + \Delta y = 2 (\sigma_x + \sigma_y) = 4 \sigma = 2 \sqrt{2} \sigma_z = 2,83 \sigma_z$.

Nous supposons que x et y sont des variables aléatoires gaussiennes. Dans ce cas, z est aussi une variable gaussienne d'espérance mathématique $Z = X + Y$. Or la probabilité pour que l'intervalle $(z - 2,83 \sigma_z ; z + 2,83 \sigma_z)$ contienne Z n'est pas 0,95, mais 0,995 : le niveau de confiance est donc augmenté ; ou encore, le risque de 5 % retenu pour x et y est devenu 0,05 % pour z : il a été divisé par 10 et a donc changé d'ordre de grandeur.

En revanche, l'addition *quadratique* des incertitudes absolues, pour une somme ou une différence, conserve le même niveau de confiance et donc le même risque.

Pour un produit ou un quotient, c'est donc l'addition quadratique des incertitudes relatives qui permet de déterminer une incertitude au même niveau de confiance que pour les mesures initiales.

Ces règles découlent de celles qui permettent de calculer la variance (carré de l'écart-type) d'une variable aléatoire z s'exprimant comme une combinaison linéaire et homogène de plusieurs autres x_1 , x_2 et x_3 :

Soit $z = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$; alors on montre que $\sigma_z^2 = \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2 + \alpha_3^2 \sigma_3^2$.

Appliquons cette règle au cas de $R = \frac{U}{I}$; alors $\frac{dR}{R} = \frac{dU}{U} - \frac{dI}{I}$, soit : $dR = \frac{dU}{I} - R \frac{dI}{I}$.

Considérant dU et dI comme des variables aléatoires indépendantes d'écart-types respectifs σ_U et σ_I , en appliquant au calcul de σ_R le résultat précédant, avec $\alpha_1 = \frac{1}{I}$ et $\alpha_2 = \frac{R}{I}$, on trouve :

$$\sigma_R^2 = \frac{\sigma_U^2}{I^2} + \frac{R^2}{I^2} \cdot \sigma_I^2, \text{ soit } \frac{\sigma_R^2}{R^2} = \frac{\sigma_U^2}{U^2} + \frac{\sigma_I^2}{I^2} \text{ ou } \frac{\sigma_R}{R} = \sqrt{\frac{\sigma_U^2}{U^2} + \frac{\sigma_I^2}{I^2}} \text{ (CQFD).}$$

De même, pour la moyenne arithmétique m de n mesures x_i de même écart-type σ , où tous les α_i sont égaux à $\frac{1}{n}$, on trouve ainsi : $\sigma_m^2 = \frac{1}{n^2} \sum_1^n \sigma^2 = \frac{\sigma^2}{n}$ soit : $\sigma_m = \frac{\sigma}{\sqrt{n}}$.

Nous avons souvent, dans les pages qui précèdent, utilisé ce résultat.

L'ancien calcul d'incertitude, qui conduisait à donner pour la moyenne de n mesures équivalentes, une incertitude absolue égale à l'incertitude absolue sur une des mesures, aboutissait, dès que n dépassait 4, par exemple, à des majorants beaucoup trop grands, ne correspondant à aucune probabilité raisonnable.

Annexe 7 : écart à la normalité et valeurs aberrantes

1) L'usage des coefficients de Student ou, plus simplement, le fait de prendre “ 2σ ” pour incertitude, suppose que l'échantillon dont on dispose puisse être considéré comme “tiré d'une population normale”.

Le tracé de la “droite de Henry”, par exemple, sur papier “gausso-arithmétique”, permet d'estimer “à l'œil” si l'ensemble des résultats recueillis est susceptible d'être tiré d'une “population gaussienne”. Il s'agit d'un papier qui tire son nom des deux graduations qu'il porte. La première, en abscisses, est banale (graduation arithmétique) ; la deuxième, en ordonnées, (de 0,01 à 0,99 ou de 0,001 à 0,999), est reliée à la fonction de répartition de Gauss, c'est-à-dire à la probabilité $P(x)$ pour qu'une variable aléatoire normale centrée et réduite soit inférieure à x . Ainsi, si l'on dispose de n mesures ordonnées par rangs croissants, on marque sur le graphique n points dont les coordonnées respectives sont les suivantes : en abscisses (sur l'échelle banale) la valeur de la mesure choisie ; en ordonnées, si le rang de la mesure est i , on porte la valeur $\frac{i}{n+1}$. Si l'on dispose de 9 mesures, leurs ordonnées respectives sont donc, respectivement : 0,1 ;

0,2 ; ... 0,9. Ensuite, on regarde si les n points obtenus sont assez proches d'une droite qui, si elle existe, porte évidemment le nom de droite de Henry. Si l'un des points, ou plusieurs, sont trop éloignés d'une droite dont les autres sont proches, il y a peu de chances que l'on puisse considérer l'échantillon comme tiré d'une population gaussienne.

Ce test visuel peut être rendu quantitatif.

2) Lorsque divers groupes d'élèves procèdent à la mesure d'une même grandeur (accélération de la pesanteur en un lieu donné, vitesse du son dans la salle de classe, concentration d'une même solution, distance focale d'une même lentille mince, dureté d'une eau, inductance d'une même bobine à une fréquence donnée, etc.), il arrive qu'une ou deux valeurs s'écartent sensiblement des autres. Parfois, l'erreur est manifeste (erreur d'opération par exemple) et on peut la corriger si les étapes du raisonnement (calibres utilisés, unités employées, etc.) ont bien été enregistrées ; parfois, en revanche, rien ne permet d'expliquer l'écart observé et l'on se demande s'il faut conserver une valeur dont la singularité n'est peut-être due qu'au hasard ou s'il faut la rejeter à cause d'un appareil faux ou d'une observation erronée.

Le test de l'écart à la moyenne fournit à cet égard un critère quantitatif. Supposons que l'on dispose de n mesures indépendantes et de poids égaux : après avoir calculé leur moyenne m et l'écart-type σ_{n-1} correspondant, on les ordonne par valeurs croissantes ; la plus faible est notée x_m et la plus grande x_M . Supposons que ce soit la plus grande qui pose problème. On calcule alors le rapport

$$v = \frac{x_M - m}{\sigma_{n-1}}$$
 Pour $n = 9$, par exemple, si les mesures suivent une loi gaussienne, v est en moyenne égal

à 1,49 ; il ne dépasse 2,24 que dans 5 % des cas. Aussi, dans ce cas, en prenant le risque de 5 % d'enlever à tort une valeur, choisit-on généralement d'écarter toute valeur qui s'éloigne de la moyenne de plus de $2,24 \sigma_{n-1}$. Le tableau ci-dessous donne, pour quelques valeurs de n , les valeurs v_{\max} de v qui, avec le même risque, ne doivent pas être dépassées :

n :	6	7	8	9	10	12	14	16	18
v_{\max} :	2,00	2,09	2,17	2,24	2,29	2,39	2,46	2,52	2,58

Le logiciel du CNDP “Incertitudes de mesure” pose les deux types de questions évoquées ci-dessus et permet d'y répondre instantanément.

Dans un premier temps, pour un échantillon donné, afin de savoir si l'échantillon des mesures peut être considéré comme tiré d'une population gaussienne, il trace automatiquement, non la droite de Henry mais celle de Pearson qui est d'inspiration très voisine. Comme les utilisateurs du logiciel ne sont pas des spécialistes des méthodes statistiques et qu'ils ne sont donc pas capables, *a priori*, de savoir si les points expérimentaux sont suffisamment alignés pour pouvoir accepter l'hypothèse d'une population gaussienne,

le logiciel prend lui-même la décision d'acceptation ou de rejet. Pour cela, il calcule le coefficient de corrélation des points expérimentaux et compare celui-ci à une valeur témoin qu'il possède en mémoire et qui dépend du nombre n de mesures. Si le coefficient de corrélation est inférieur à la valeur témoin, l'hypothèse de normalité est rejetée, sinon elle est acceptée. Les différents coefficients témoins (pour n compris entre 5 et 50) ont été calculés tout spécialement pour ce logiciel selon la méthode dite de Monte Carlo. Ils s'échelonnent entre $r = 0,8752$ pour $n = 5$ et $r = 0,9760$ pour $n = 50$ en passant par $r = 0,9186$ pour $n = 10$ et $r = 0,9496$ pour $n = 20$.

L'hypothèse de non normalité peut être rejetée pour plusieurs raisons :

- 1) parce que la population mère n'est pas normale (ou gaussienne) ;
- 2) parce que l'échantillon contient des valeurs aberrantes.

Le logiciel, en effectuant pour chaque mesure le test de l'écart à la moyenne exposé ci-dessus, propose donc d'éliminer celles des valeurs qui peuvent, avec une bonne probabilité, être considérées comme aberrantes.

Si l'opérateur accepte de les ôter de l'échantillon, il reprend sa démarche et cherche à savoir si les mesures conservées peuvent à nouveau être considérées comme tirées d'une population gaussienne, ce qui est très généralement le cas.

Il en découle une grande simplicité d'utilisation car les élèves, libérés de calculs fastidieux peuvent davantage se consacrer aux problèmes inhérents aux mesures. Informés par l'ordinateur de la présence de mesures aberrantes, ils acceptent plus volontiers la remise en cause de telle ou telle de leurs mesures.

Annexe 8 : comparaison des écarts-types de deux échantillons

Il peut arriver qu'un professeur fixe pour but d'un TP la comparaison de deux méthodes de mesures ; ou encore, la même mesure ayant été proposée aux "binômes" de chacun des deux groupes d'une même classe, il peut se poser la question de savoir s'il peut regrouper l'ensemble des résultats de la classe afin de disposer d'un plus grand nombre de mesures indépendantes dans le but d'améliorer la précision de la mesure collective.

Dans chacun de ces deux cas, il faut comparer les écart-types relatifs aux deux échantillons de mesure.

Notons n_1 l'effectif du premier échantillon (le nombre de mesures indépendantes d'une même grandeur effectuées par les élèves en utilisant la méthode n°1 dans le premier cas envisagé) et n_2 l'effectif du deuxième. C'est la loi $F(v_1, v_2)$ dite de Fisher qui permet de comparer les écarts-types (ou plutôt les variances de ces deux échantillons). Les nombres v_1 et v_2 sont respectivement égaux à $(n_1 - 1)$ et $(n_2 - 1)$, ce sont les degrés de liberté respectifs des variances considérées. Ainsi, lorsque l'effectif n de chaque échantillon est égal à 9, les nombres v_1 et v_2 de degrés de liberté des variances $(\sigma_{1\ n-1})^2$ ou $(\sigma_{2\ n-1})^2$ sont égaux à $n-1 = 8$.

La loi de Fisher donne la distribution du rapport de $(\sigma_{1\ n-1})^2 / (\sigma_{2\ n-1})^2$ dans le cas où les deux échantillons sont tirés de la même population mère gaussienne (c'est ce que l'on appelle faire l'hypothèse nulle). Il permet de connaître la probabilité pour que, dans le cadre de cette hypothèse, telle ou telle valeur de ce rapport $F(v_1, v_2)$ soit dépassée du seul fait du hasard. Pour deux échantillons (1) et (2) de 9 valeurs tirées d'une même population gaussienne, par exemple, le rapport $F(v_1, v_2)$ est supérieur à 4,43 dans 2,5 % des cas. Mais, en inversant numérateur et dénominateur, le rapport $F(v_2, v_1)$ est aussi supérieur à 4,43 dans 2,5 % des cas. Au total, le rapport de la plus grande des variances à la plus petite est supérieur à 4,43 dans 5 % des cas.

C'est le critère que nous retenons, dans ces conditions ($n_1 = n_2 = 9$), pour considérer comme significativement différentes deux variances. Comme le rapport correspondant des écarts-types $(\sigma_{1n-1} / \sigma_{2n-1})$ ou l'inverse $(\sigma_{2n-1} / \sigma_{1n-1})$, est égal à la racine carrée de 4,43 soit 2,1, nous admettrons ceci :

Dans le cas de deux échantillons de 9 mesures d'une même grandeur, si l'un des rapports $(\sigma_{1n-1} / \sigma_{2n-1})$ ou $(\sigma_{2n-1} / \sigma_{1n-1})$ est supérieur à 2,1, alors nous ne pouvons pas considérer que les deux échantillons sont tirés de la même population gaussienne (le risque de rejeter cette hypothèse à tort est ainsi égal à 5 %).

Inversement, nous traiterons (cf. plus loin) le cas de la mesure de la distance focale d'une lentille mince selon deux méthodes : la méthode de Silberman ayant conduit à 9 mesures d'écart-type $\sigma_{S_{n-1}} = 0,163$ cm et la méthode de l'autocollimation ayant, pour 9 mesures également, conduit à $\sigma_{a_{n-1}} = 0,101$ cm. Le rapport $\sigma_{S_{n-1}} / \sigma_{a_{n-1}}$ étant égal à 1,61, valeur nettement inférieure à 2,1, on ne peut pas, par ces seules mesures, considérer que la méthode de Silberman est significativement moins précise que la méthode d'autocollimation.

Exemples de traitement de mesures en travaux pratiques ²

Des explications complémentaires, destinées à l'information des seuls professeurs, sont insérées en italiques.

Chimie : détermination d'une grandeur par dosage

Contexte de l'étude

L'exemple a été choisi en 1^{ère} S dans le cadre de l'option ; il s'agit de déterminer par complexométrie la dureté d'une eau, c'est-à-dire sa concentration totale en ions calcium et magnésium ; l'étiquette de l'eau de Contrexéville étudiée indique 486 mg/L d'ions Ca^{2+} et 84 mg/L d'ions Mg^{2+} .

Les élèves répartis en deux groupes de neuf binômes ont utilisé un matériel classique de dosage : burette de 25,0 mL, pipette jaugée à deux traits de 10,0 mL et agitateur magnétique. Lors de cette séance, ils ont réalisé leur second dosage, le premier étant le dosage des ions chlorure par la méthode de Mohr.

Un point à signaler lors du déroulement de la séance : les élèves du premier groupe s'étant trop précipités en début de séance (rinçages mal gérés, prédosage parfois absent...), des consignes plus explicites ont été données aux élèves du second groupe qui ont alors tous manipulé avec beaucoup de soin.

Résultats bruts

Le dosage consiste à déterminer le volume équivalent V_E de solution d'Edta qu'il faut verser pour complexer la totalité des ions calcium et magnésium initialement présents dans l'échantillon d'eau de volume V ; connaissant la concentration $[\text{Y}^{4-}]$ de l'Edta en solution, on calcule la concentration totale C des ions calcium et magnésium dans l'eau avec la relation : $C = \frac{[\text{Y}^{4-}] \cdot V_E}{V}$.

On aboutit alors au calcul de la dureté de l'eau exprimée en degrés hydrotimétriques français notée D et définie par la relation : $D = 10 \times C$, les concentrations molaires des ions calcium et magnésium étant exprimées en mmol.L^{-1} ; on exprime alors D en °TH.

(N.B : une eau est considérée dure lorsque D est supérieure à 50).

Il faut noter que les conditions expérimentales ne permettent d'exprimer le résultat qu'avec trois chiffres significatifs ; en effet V_E n'est connu qu'à 0,1 mL près (exemple de résultat : $V_E = 15,6$ mL).

Résultats ordonnés du premier groupe, exprimés en °TH :

153, 154, 154, 155, 155, 156, 156, 160, 165.

Moyenne : 156,4 °TH ; estimateur de l'écart-type, noté σ_{n-1} : 3,8 °TH.

Résultats ordonnés du deuxième groupe, exprimés en °TH :

154, 155, 155, 155, 156, 156, 156, 157.

Moyenne : 155,6 °TH ; estimateur de l'écart-type : σ_{n-1} : 0,88 °TH.

² Les TP ont été réalisés dans la classe de Mme Colette Huet, du Lycée David d'Angers

Analyse des résultats en classe

L'indication de l'étiquette permet de calculer le degré hydrotimétrique de référence ; en désignant par x la concentration massique en ions Ca^{2+} et par y la concentration massique en ions Mg^{2+} , x et y étant exprimées en $\text{mg}\cdot\text{L}^{-1}$, on a :

$$D = 10 \times \left(\frac{x}{40,08} + \frac{y}{24,31} \right) ; \text{ soit : } D = 155,8 \text{ } ^\circ\text{TH}.$$

Individuellement, chacun des binômes met en relation son résultat avec la valeur de référence en évaluant par exemple l'écart relatif entre les deux valeurs. Mais ceci est très pauvre !

La mise en commun des résultats des neuf binômes du même groupe permet, en revanche, d'aborder le traitement statistique d'une série de mesures.

Pour chacun des groupes, la moyenne, estimateur de la grandeur mesurée, est en très bon accord avec l'indication de l'étiquette. Les deux groupes ont à ce niveau des résultats très proches (le même résultat, si on se limite à trois chiffres significatifs).

Avec une étendue de 12 °TH dans le premier cas et 3 °TH dans le second, ou un écart-type 4 fois plus grand pour le premier groupe que pour le second, il est évident que les résultats globaux du premier groupe sont beaucoup plus dispersés que ceux du second. Le second groupe a fourni un travail expérimental de bien meilleure qualité : la seule moyenne des résultats ne le montre pas ; l'impact de la qualité du travail expérimental (rinçages, prédosage...) sur le résultat est ainsi parfaitement mis en valeur.

Commentaires et approfondissements (pour les professeurs)

L'objectif de ce paragraphe est d'utiliser la situation expérimentale décrite ci-dessus pour aller plus loin en ce qui concerne la résolution de différents problèmes tels que :

Combien de chiffres significatifs est-il raisonnable de garder dans un résultat ?

Comment exprimer un résultat pour qu'il prenne en compte la précision de la mesure ?

Comment reconnaître une mesure aberrante ?

Il s'agit aussi de concrétiser certains points théoriques exposés dans la première partie de ce chapitre.

- *Calcul du degré hydrotimétrique de référence D*

Le calcul aboutit à $D = 155,8 \text{ } ^\circ\text{TH}$; faut-il garder les 4 chiffres significatifs ?

Malheureusement, on ne connaît pas les incertitudes relatives aux concentrations x et y , respectivement en ions Ca^{2+} et Mg^{2+} , ce qui rend impossible la recherche d'un intervalle de confiance à partir de l'étiquette et qui pose également un problème pour l'expression de D car on ne sait exactement combien de chiffres significatifs on doit conserver.

Puisque les expressions numériques de x et de y sont arrondies (milligramme par litre), supposons qu'au niveau de confiance 95 %, on ait : $\Delta x = \Delta y = 0,5 \text{ mg}\cdot\text{L}^{-1}$.

Alors, au même niveau de confiance, l'incertitude ΔD est :

$$\Delta D = \sqrt{\left(\frac{10 \times \Delta x}{40,08} \right)^2 + \left(\frac{10 \times \Delta y}{24,31} \right)^2} = 0,24 \text{ } ^\circ\text{TH}.$$

Ceci permet d'écrire $D = (155,8 \pm 0,24) \text{ } ^\circ\text{TH}$.

Il n'est pas bon, en effet, d'arrondir D à 156 °TH, car alors on commettrait une erreur d'écriture presque égale à l'incertitude ΔD . Pour renoncer au quatrième chiffre significatif et écrire $D = 156 \text{ } ^\circ\text{TH}$, ce qui représente une translation de 0,2 °TH, il faudrait avoir $\Delta D > 1 \text{ } ^\circ\text{TH}$ car nous nous interdisons des translations supérieures à 0,2 ΔD . En les supposant égales, les incertitudes Δx et Δy devraient alors être supérieures à 2 $\text{mg}\cdot\text{L}^{-1}$, ce qui est certainement exagéré (les concentrations données par la marque Perrier sont arrondies à 0,1 $\text{mg}\cdot\text{L}^{-1}$). Nous conservons donc l'écriture : $D = 155,8 \text{ } ^\circ\text{TH}$.

- *Comparaison des mesures des deux groupes*

Du fait des consignes supplémentaires et de leur impact évident, on ne peut considérer que les mesures des deux groupes constituent deux échantillons tirés d'une même population : le "poids" d'une mesure, selon qu'elle émane d'un groupe ou de l'autre, n'est pas le même. Les résultats de chaque groupe doivent

donc être analysés séparément (*nous avons vu en annexe 7 qu'un rapport des écarts-types égal à 2,1 est généralement considéré comme suffisant pour marquer la dissemblance de deux échantillons de neuf mesures*).

• *Analyse des résultats du second groupe*

On constate que l'accord entre la moyenne des mesures du second groupe (155,6 °TH) et la valeur calculée d'après l'étiquette (155,8 °TH) est excellent.

Notons ΔD la demi-largeur de l'intervalle de confiance de D, au niveau de confiance 95 %, calculée d'après les résultats de ce second groupe :

$$\Delta D = \frac{t \cdot \sigma_{n-1}}{\sqrt{9}}, \text{ avec } t, \text{ coefficient de Student tel que } t = 2,31, \text{ soit : } \Delta D = \frac{2,31 \times 0,88}{3} = 0,68 \text{ °TH}$$

• *Expression définitive du résultat du second groupe*

À cause de l'imprécision sur la mesure de V_E , les neuf mesures individuelles, exprimées à l'aide de trois chiffres significatifs seulement, ne correspondent en fait qu'à 4 valeurs.

De ce fait, il serait illusoire d'exprimer la moyenne (155,56 °TH) avec cinq chiffres significatifs. Nous partons donc de $D = (155,6 \pm 0,68) \text{ °TH}$.

L'intuition nous conduit à écrire $D = (155,6 \pm 0,7) \text{ °TH}$. C'est bien ce que confirment les critères que nous utilisons et qui sont basés sur le fait que le niveau de confiance choisi (ici 95 %) ne doit pas être affecté de plus de 1 % de variation par chaque opération de simplification à laquelle on se livre.

Critères retenus :

- On peut supprimer des chiffres significatifs dans l'expression de l'incertitude ΔD (et procéder aux arrondis correspondants), à condition que ces opérations provoquent une variation de l'incertitude ΔD initiale inférieure à 4 %. Ici $4 \times \frac{\Delta D}{100} = 0,03 \text{ °TH}$; on peut donc remplacer 0,68 °TH par 0,7 °TH, ce qui entraîne seulement une variation de 0,02 °TH.

- On peut supprimer des chiffres significatifs sur D (et procéder aux arrondis correspondants) à condition que ces opérations ne provoquent pas un décalage de D supérieur à 0,2 ΔD . Ici, $0,2 \Delta D = 0,14 \text{ °TH}$; le passage de 155,6 à 156 °TH entraînerait un décalage de 0,4 °TH trop important : cette erreur d'écriture, supérieure à la moitié de l'incertitude, abaisserait exagérément le niveau de confiance (*qui passerait à 83 %*).

• *Analyse des résultats du premier groupe*

La plus grande des mesures (165 °TH) est vraiment très éloignée des autres : son écart à la moyenne est égal à 2,26 σ_{n-1} . Nous remarquons que, du fait de cette mesure, certainement due à une manipulation effectuée sans soin, la moyenne (156,4 °TH) des résultats du premier groupe d'élèves n'appartient même pas à l'intervalle de confiance au niveau de confiance 95 % du groupe 2, soit [149,9 °TH ; 156,3 °TH], alors qu'en ôtant cette mesure, la nouvelle moyenne (155,4 °TH) en fait pleinement partie (*cette dernière constatation ne constitue cependant pas un critère de rejet*).

Divers tests qualitatifs ou quantitatifs permettent de décider de rejeter la valeur 165 °TH : le tracé de la droite de Pearson, grâce au logiciel "Incertitudes de mesure" du CNDP, montre que, compte tenu de cette mesure, l'échantillon étudié ne peut être considéré comme tiré d'une population gaussienne ; le test quantitatif de l'écart à la moyenne, lui, recommande d'enlever celle qui s'écarte de la moyenne de plus de 2,24 σ_{n-1} : c'est le cas de la valeur considérée, éloignée de la moyenne de 2,26 σ_{n-1} .

Il est donc plus sage de ne pas tenir compte de cette mesure.

Pour le groupe 1, les mesures conservées sont donc, ordonnées et exprimées en °TH :

153, 154, 154, 155, 155, 156, 156, 160.

Moyenne : 155,4 °TH ; estimateur de l'écart-type, noté σ_{n-1} : 2,1 °TH.

L'accord entre cette nouvelle moyenne et la valeur calculée d'après l'étiquette est encore très bon, mais

les écarts-types des deux séries de mesures sont toujours beaucoup trop différents pour que celles-ci puissent être considérées comme de poids égaux.

- *Expression définitive du résultat du premier groupe*

La demi-largeur ΔD de l'intervalle de confiance, au niveau de confiance 95 %, calculée d'après les mesures retenues, avec $t = 2,37$, puisque désormais $n = 8$, donne : $D = 1,76 \text{ }^\circ\text{TH}$, quantité qui peut être arrondie à $1,8 \text{ }^\circ\text{TH}$ (la variation correspondante de ΔD est inférieure à 4 %).

Le résultat définitif du premier groupe est donc : $D = (155,4 \pm 1,8) \text{ }^\circ\text{TH}$: on remarque en effet qu'en arrondissant la moyenne M des huit mesures de ce groupe à $155 \text{ }^\circ\text{TH}$, on provoquerait un décalage supérieur à $0,2 \Delta D = 0,36 \text{ }^\circ\text{TH}$.

Dans ce cas, l'application de la loi de Student montre que le passage de $D = (155,4 \pm 1,76) \text{ }^\circ\text{TH}$ à $D = (155 \pm 1,8) \text{ }^\circ\text{TH}$ entraîne une variation du niveau de confiance de 95 % à 93,8 % car les deux arrondis ont des effets contraires : le passage de $\Delta D = 1,76 \text{ }^\circ\text{TH}$ à $D = 1,8 \text{ }^\circ\text{TH}$ a pour effet d'augmenter le niveau de confiance (de 95,0 % à 95,4 %) ; le passage de $D = 155,4 \text{ }^\circ\text{TH}$ à $D = 155 \text{ }^\circ\text{TH}$ ramène le niveau de confiance à 93,8 %, ce qui, après tout, peut être accepté par certains comme suffisamment voisin de 95 %.

Le résultat collectif du premier groupe, $D = (155,4 \pm 1,8) \text{ }^\circ\text{TH}$ est malgré tout un résultat dont la précision relative, $\Delta D / D$, égale à 1,2 %, doit être considérée comme bonne pour des élèves réalisant leur deuxième dosage.

Optique, comparaison de deux méthodes de détermination d'une distance focale

Contexte de l'étude

Les résultats ci-dessous ont été obtenus, pour une même lentille, par un groupe d'élèves de terminale S en spécialité. Les élèves ont utilisé un matériel classique de lycée : banc d'optique gradué à 0,2 cm ; supports de lentille et d'écran avec repères.

Résultats bruts

Mesures de distances focales, exprimées en centimètres et ordonnées :

- Méthode de Silberman (la distance focale est égale au quart de la distance mesurée, distance pour laquelle l'image et l'objet sont de même taille) : 4,65 ; 4,73 ; 4,75 ; 4,75 ; 4,80 ; 4,88 ; 4,90 ; 4,95 ; 5,20.

Moyenne : $m_S = 4,85 \text{ cm}$; estimateur de l'écart-type : $\sigma_{S, n-1} = 0,163 \text{ cm}$.

- Méthode d'autocollimation (la distance focale est directement mesurée) :

4,8 ; 4,8 ; 4,9 ; 4,9 ; 5,0 ; 5,0 ; 5,0 ; 5,0 ; 5,1.

Moyenne : $m_a = 4,94 \text{ cm}$; estimateur de l'écart-type : $\sigma_{a, n-1} = 0,101 \text{ cm}$.

Analyse des résultats en classe

Les deux méthodes donnent des moyennes concordantes.

Les résultats obtenus avec la méthode d'autocollimation sont moins dispersés (étendue et écarts-types plus faibles) ; cette méthode semble plus performante que la méthode de Silberman.

Commentaires et approfondissements

- *Valeurs aberrantes ; précision du matériel*

Dans les mesures relatives à la méthode de Silberman, la distance focale de 5,20 cm semble un peu écartée des autres. En fait, elle ne diffère de la moyenne 4,85 cm que $2,15 \sigma_{S, n-1}$ ce qui n'est pas suffisant pour la rejeter. Aucune des autres mesures n'est suspecte.

Le matériel (graduation du banc, justesse des repères...) n'est pas assez précis pour que cette étude soit totalement convaincante puisque huit valeurs pour la méthode de Silberman et quatre pour la méthode d'autocollimation suffisent à exprimer neuf mesures.

- *Expression des résultats*

- Méthode de Silberman : calcul de l'incertitude, considérée comme la demi-largeur de l'intervalle de confiance au niveau de confiance 95 % :

$$\Delta f = \left(2,31 \times \frac{0,163}{3} \right) \text{ cm} = 0,126 \text{ cm.}$$

Nous pouvons arrondir Δf à 0,13 cm sans modifier sensiblement le niveau de confiance affiché.

Peut-on enlever un chiffre significatif à la moyenne $m_S = 4,85$ cm et estimer f par 4,9 cm ?

La translation qui en résulterait serait de 0,05 cm, représentant 0,40 Δf ; elle est trop grande. On a donc :

$$f_S = (4,85 \pm 0,13) \text{ cm}$$

- Méthode d'autocollimation : au niveau de confiance 95 %, $\Delta f = 0,078$ cm, quantité que nous pouvons confondre avec 0,08 cm.

La translation qui résulterait du passage de 4,94 cm à 4,9 cm serait là encore de 0,04 cm soit 0,51 Δf , ce qui, *a fortiori*, n'est pas acceptable. On a donc :

$$f_a = (4,94 \pm 0,08) \text{ cm}$$

- *Comparaison des estimateurs f_S et f_a de la distance focale f*

Les deux intervalles de confiance ont une partie commune (4,86 cm ; 4,99 cm), ce qui montre que les résultats donnés par les deux méthodes sont compatibles. Mais on peut se demander si le fait que l'estimateur f_S donné par la méthode de Silberman soit juste à la limite de l'intervalle de confiance déduit de la méthode d'autocollimation, n'est pas le signe que des erreurs systématiques ont été commises. Au lycée, par exemple, l'index d'une lentille ne correspond pas toujours exactement à la position de son centre optique. Ainsi, dans la méthode d'autocollimation par exemple, ne pas retourner une lentille par rapport au sens de propagation de la lumière lorsque celle-ci passe d'un groupe à l'autre, peut entraîner une erreur systématique de plusieurs millimètres, la retourner peut donner des résultats dispersés.

Compte tenu des valeurs de $\sigma_{S_{n-1}}$ et de $\sigma_{a_{n-1}}$ la différence $f_S - f_a$ n'est cependant pas significative.

- *Comparaison des écarts-types*

Le rapport des écarts-types est de 1,61. Il semble indiquer que, toutes causes d'erreurs confondues, la mise en œuvre de la méthode d'autocollimation par les élèves de terminale S, avec le matériel limité qui est mis à leur disposition (absence de viseurs) donne des résultats plus précis que celle de la méthode de Silberman.

Cette valeur n'est cependant pas suffisamment élevée pour poser une telle assertion (annexe 8).

On pourrait en effet montrer que, même si les deux méthodes étaient de précision tout à fait identiques, la valeur 1,61 caractérisant le rapport des estimateurs $\sigma_{S_{n-1}}$ et de $\sigma_{a_{n-1}}$ (quel que soit celui que l'on place au dénominateur) serait naturellement dépassée dans 18 % des cas.

On retiendra donc de ce TP d'optique que la méthode expérimentale utilisée pour comparer deux méthodes de détermination de la distance focale d'une lentille n'a pas permis, ici, avec les protocoles utilisés, de les classer, le rapport du plus grand des estimateurs ($\sigma_{S_{n-1}}$) au plus petit ($\sigma_{a_{n-1}}$), n'étant pas assez élevé (1,61).