

Quelques notions élémentaires et incomplètes à propos des incertitudes de mesure

Introduction : éléments de base

Dans la mise en oeuvre de l'étude rationnelle des phénomènes naturels, on est amené à définir des *grandeurs physiques*. En première approche, on appelle « grandeur physique » toute propriété rattachée à l'étude d'un phénomène et qui peut s'évaluer, se quantifier, au travers d'un nombre.

Appréhender un phénomène, c'est en partie - *et en partie seulement* - savoir distinguer ou inventer les grandeurs physiques qui lui sont rattachées, puis découvrir par la théorie ou par l'expérience les liens qui les relient entre elles.

On a reconnu depuis longtemps qu'il existe des propriétés - et donc des grandeurs physiques - de natures *a priori* différentes : la distance parcourue, la durée qui s'écoule, la masse d'un objet,..., en sont des exemples rattachés à la Physique. Chacune de ces grandeurs est généralement affectée d'une *dimension* (une longueur : [L] ; un temps : [T],...) qui permet de reconnaître sa nature.

On définit pour chaque grandeur physique la valeur *étalon* de cette grandeur que l'on choisit égale à l'*unité* (par exemple, le mètre (*m*) pour la dimension longueur). Le nombre qui permet de quantifier une grandeur physique est alors compris comme : le nombre de fois où l'on retrouve la valeur de l'étalon dans la grandeur physique en question.

Le Trésor de la Langue Française Informatisé définit la *mesure* comme l'évaluation d'une grandeur physique ou d'une quantité, par comparaison avec une autre de même espèce, prise comme terme de référence (i.e. l'étalon).

La mesure d'une grandeur physique est toujours accompagnée d'*erreurs de mesure* aux origines diverses que l'on a l'habitude de séparer en deux classes non forcément disjointes.

Il peut s'agir d'*erreurs systématiques* comme par exemple l'utilisation d'un étalon de mesure faussé : une règle ayant un défaut de graduation à la construction, une balance mal tarée, un composant électronique défectueux au sein de l'appareil de mesure, un défaut d'isolation électrique... Ce genre d'erreur s'élimine autant que possible par le contrôle de l'appareil de mesure, de son étalonnage et de ses conditions d'utilisation, mais aussi et surtout par l'emploi successif de méthodes et d'appareils différents.

Les *erreurs aléatoires* ou *accidentelles* sont généralement imputables à l'imperfection des sens de celui ou celle qui réalise la mesure, ou à l'instabilité de l'appareil de mesure au cours de l'expérience : sensibilité aux variations de température, à la tension d'alimentation électrique, aux vibrations, aux perturbations électromagnétiques des appareils environnants, etc... On ne peut que tenter de limiter au maximum ce genre d'erreur par un contrôle soigné des phénomènes perturbateurs sans pouvoir, dans la pratique, les éliminer totalement. La mesure répétée d'une même grandeur physique avec le même appareil fournit donc dans la pratique des valeurs légèrement différentes à chaque fois, visibles ou non selon la sensibilité de l'appareil.

L'appareil de mesure lui-même est une source d'imprécision : une règle graduée (étalonnée) au mm ne pourra naturellement pas permettre de distinguer le centième de mm ; une personne à l'œil exercé pourra peut-être évaluer le dixième de mm avec cette règle, mais on admet conventionnellement que la précision de la mesure ne sera pas meilleure que la demi-distance séparant deux graduations.

De même, les différents éléments qui composent un voltmètre (numérique ou analogique) sont construits à partir de matériaux dont la sensibilité à des variations de tension ou de courant n'est bien connue que sur une certaine plage de valeurs (par exemple 200 - 1000 Volts) et reste usuellement limitée à quelques pour milles de ces valeurs.

L'origine des erreurs de mesure aléatoire est donc surtout liée au contrôle imparfait des conditions expérimentales dans lesquelles s'effectue la mesure ainsi qu'à la fabrication imparfaite des appareils de mesures dont la sensibilité est limitée.

On peut bien sûr construire des appareils de mesure de plus en plus précis ; mais alors on devient de plus en plus sensible aux conditions dans lesquelles s'effectue la mesure.

En outre, plus on recherche une grande précision de mesure, plus on se heurte au problème épistémologique de savoir jusqu'à quel point une grandeur physique peut-elle s'identifier à un nombre précis ? Est-il de bon sens d'affirmer, par exemple, que la longueur exacte d'une table est de $2,23456789123456789... m$? Il y a là une difficulté de principe : selon la théorie atomique, la table est constituée d'atomes et il en est de même pour l'appareil utilisé pour la mesure. Comment alors définir la longueur de la table et celle de l'étalon avec une précision meilleure que celle de la taille d'un atome (environ 1 \AA , soit $10^{-10} m$) ? Saura-t-on définir la taille précise d'un atome ? Saura-t-on définir l'atome qui constitue le bout de la table ? ...

Vouloir une précision de plus en plus fine, c'est donc se heurter, en dernier ressort, à la question de la signification théorique et pratique de la grandeur physique à mesurer.

Pour toutes les raisons qui viennent d'être évoquées, on est conduit à encadrer entre deux valeurs extrêmes la valeur d'une grandeur physique mesurée. La grandeur physique, supposée posséder une valeur intrinsèque bien déterminée, ne sera donc connue qu'avec une certaine *incertitude*.

Les différents traitements qui vont suivre à propos de ces incertitudes de mesure concernent surtout les erreurs aléatoires. Les erreurs systématiques les plus flagrantes sont supposées avoir été éliminées : les appareils de mesure fonctionnent correctement dans les conditions indiquées par le constructeur, et l'expérimentatrice est en pleine possession de ses sens.

Même dans ce cadre restreint, l'estimation des erreurs de mesure reste une science approximative au sens où cette estimation fournira seulement l'ordre de grandeur de l'erreur commise. Dans ces conditions, on s'attachera à utiliser plutôt des méthodes qui fournissent une erreur possible maximale sans toutefois la surestimer abusivement. Enfin, il faut insister sur le fait que les méthodes utilisées dépendent grandement des conditions mises en place pour la mesure et de la nature de la grandeur physique mesurée.

N.B. Ce polycopié est librement inspiré des nombreux autres polycopiés en téléchargement libre sur le net ainsi que de sites web, et/ou de plusieurs livres et autres polycopiés qui traînent dans les bibliothèques et qui traitent tous du sujet des incertitudes de mesure ainsi que des méthodes statistiques d'analyses de données. Il est vivement conseillé aux étudiant(e)s intéressé(e)s par la matière de s'emparer de toute cette culture gratuite ; non pas pour la consommer mais pour se la ré-approprier, la penser, la critiquer. Le conseil est naturellement valable pour tout autre sujet.

Vocabulaire, définitions et conventions classiques

Soit G une grandeur physique supposée posséder une valeur bien déterminée que l'on appellera sa valeur exacte ou vraie : g_e .

Lors de la mesure de G on n'a aucune raison d'affirmer que le nombre obtenu exprime sa valeur exacte puisqu'une nouvelle mesure donne généralement lieu à un nombre légèrement différent. Supposons alors que nous ayons réalisé n mesures de G qui ont fourni les résultats g_1, g_2, \dots, g_n . La valeur moyenne expérimentale (ou empirique), g , de ces résultats est :

$$g = \frac{1}{n} \sum_{i=1}^n g_i,$$

et elle est appelée : valeur approchée ou mesurée de G .

Il arrive assez souvent qu'une seule mesure de G ait été réalisée parce que, par exemple, l'appareil fournit toujours la même valeur si on répète la mesure : cela signifie que la sensibilité de cet appareil est telle qu'une seule mesure a suffi pour obtenir g .

Si la mesure est unique parce que le temps pris pour réaliser une telle mesure est très long, alors on accepte que la valeur g_1 soit prise pour g .

On appelle erreur absolue sur la mesure de G la différence : $\delta g \equiv |g - g_e|$. Cette quantité n'est pas connue puisqu'en général on ne connaît pas g_e ¹. Aussi devra-t-on se contenter d'en estimer une limite supérieure, Δg , appelée incertitude absolue. C'est cette quantité que l'on s'attache à déterminer en tenant compte, dans la mesure du possible, de toutes les circonstances pouvant avoir une influence non négligeable sur sa valeur.

Par définition, Δg est un nombre positif. Une fois ce nombre connu, on écrit le résultat de la mesure de G sous la forme :

$$G = (g \pm \Delta g) \text{ unités}$$

Cette écriture doit garantir en principe que la valeur exacte de G est comprise dans l'intervalle : $g - \Delta g \leq g_e \leq g + \Delta g$. Et l'on dira de G que sa valeur exacte est connue à (moins de) Δg unités près. Par exemple : la longueur de la table est $L = (\ell \pm \Delta \ell) m = (2.23 \pm 0.02) m$; sa valeur exacte est donc connue à 2 cm près.

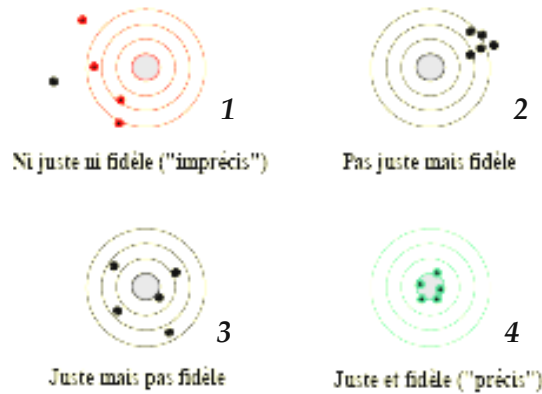
La façon d'écrire le résultat de la mesure doit tenir compte de la valeur de Δg : si l'on a trouvé après n mesures que la valeur approchée (ℓ) de la longueur de la table est 2.2345...m alors que $\Delta \ell$ vaut 0.02 m, cela entraîne que tous les chiffres caractérisant ℓ au-delà des centièmes de m ne sont pas significatifs et n'ont pas lieu d'être retenus dans le résultat final.

Comment estime-t-on Δg ? En général, le constructeur de l'appareil de mesure a réalisé des étalonnages de l'appareil, en ce sens qu'il a mesuré différents étalons avec cet appareil, ce qui

¹ On ne connaît g_e que si la grandeur physique à mesurer a été déclarée comme étalon ; par exemple, la vitesse de la lumière dans le vide.

lui a permis de le régler et d'estimer la valeur de Δg . Cette valeur est donc fournie avec la notice de l'appareil.

Si, pour quelque raison que ce soit, on ne connaît pas le Δg de l'appareil utilisé, il peut s'avérer nécessaire de l'étalonner. Si l'étalonnage est satisfaisant, on peut réaliser n mesures de la grandeur physique, sélectionner les valeurs minimale, g_{min} , et maximale, g_{max} , obtenues et prendre pour Δg la différence positive de ces valeurs : $\Delta g = g_{max} - g_{min}$, dont on ne conserve en général que le premier chiffre non nul, compte tenu du fait que Δg doit représenter une limite supérieure (par exemple, si on trouve par ce moyen $g_{max} - g_{min} \cong 0.015 m$, certains prendrons : $\Delta g = 0.02 m$).



Métaphore de la cible : la valeur exacte, g_e , de la grandeur physique G est représentée par le centre de la cible et les « coups » sur la cible représentent les différentes valeurs g_1, g_2, \dots, g_5 obtenues après 5 mesures. **En 1 :** il y a une erreur systématique et une grande erreur aléatoire (incertitude absolue importante). **En 2 :** erreur systématique toujours présente mais incertitude absolue faible. **En 3 :** erreur systématique éliminée mais incertitude absolue importante. **En 4 :** erreur systématique éliminée et faible incertitude absolue

On se rend mieux compte de la précision d'une mesure en calculant le rapport de l'incertitude absolue de la mesure à sa valeur approchée : $\Delta g/g$; on appelle ce rapport : incertitude relative de la mesure. On l'exprime généralement en pourcent, pour mille, etc...

Il est par exemple trois fois plus précis d'avoir mesuré - à un jour et une heure donnés ² - la distance Terre-Lune à $1000 km$ près que d'avoir mesuré la longueur de notre table à $2 cm$ près. En effet, pour la distance Terre-Lune l'incertitude relative sera de : $1000/385000 \cong 0.0026$, soit environ 3 ‰ ; tandis que pour notre table on obtient : $0.02/2.23 \cong 0.0089$, soit 9 ‰ .

En général, le constructeur d'un appareil de mesure fournit la précision de l'appareil sous la forme d'une incertitude relative.

On définit aussi l'erreur relative de la mesure par le rapport : $\delta g/g_e$. Cette notion entre en jeu dès lors que la grandeur physique mesurée est un étalon, car alors on connaît g_e .

N.B. Dans le langage courant, on confond souvent les deux mots : erreur et incertitude

² Il est important de préciser le jour et l'heure de la mesure ici car la distance Terre-Lune varie entre $360\,000 km$ et $409\,000 km$ du fait de la trajectoire elliptique de la Lune autour de la Terre. Valeur moyenne : $385\,000 km$.

Propagation des erreurs

Il n'est pas rare qu'une fois la grandeur physique G mesurée, on ait besoin de calculer le double de cette grandeur, ou son carré, ou toute autre opération mathématique. Comme G n'est connue qu'à Δg près, la question se pose alors d'estimer l'incertitude absolue du résultat obtenu : comment se propage l'incertitude initiale dans le calcul à effectuer ?

Le bon sens commande de remplacer G par ses valeurs encadrées, $g \pm \Delta g$, dans ce calcul. Soit par exemple à estimer G^3 ; on aura ainsi :

$$G^3 = (g \pm \Delta g)^3 \text{ unités}^3 = [g^3 \pm 3g^2\Delta g + 3g(\Delta g)^2 \pm (\Delta g)^3] \text{ unités}^3$$

Pour notre règle, on obtient le résultat :

$$L^3 = (\ell \pm \Delta \ell)^3 m^3 = (2.23 \pm 0.02)^3 m^3 = (11.089567 \pm 0.298374 + 0.002676 \pm 0.000008) m^3$$

où il est manifeste que les deux derniers termes de cette somme algébrique peuvent être négligés au regard de l'incertitude engendrée par le terme proportionnel à Δg . Au final, la grandeur physique « volume », V , calculée à partir de la mesure de la longueur de la table, L , s'écrira :

$$V = (v \pm \Delta v) m^3 = (11.1 \pm 0.3) m^3$$

La valeur approchée, v , de la grandeur physique ainsi calculée est donc égale au cube de la valeur approchée de la longueur de la table, ℓ , tandis que son incertitude absolue vaut : $\Delta v = 3\ell^2\Delta\ell$.

Cet exemple montre que l'estimation de l'incertitude absolue sur le résultat d'un calcul réalisé à partir d'une grandeur physique mesurée se confond avec le *calcul différentiel*. On adoptera donc le point de vue général suivant :

Soit X une grandeur physique dont la mesure a fourni : $X = (x \pm \Delta x) \text{ unités}$; et soit Y la grandeur physique construite à partir de X telle que : $Y = f(X)$ où f est une fonction mathématique connue. Une variation dX de X entraîne une variation dY de Y qui s'écrit, selon les règles usuelles du calcul différentiel, et à l'ordre 1 en dX ³ :

$$dY = f'(X) dX$$

où f' est la fonction dérivée de la fonction f .

La valeur de Y déduite de la mesure de X sera alors :

$$Y = (y \pm \Delta y) \text{ unités}$$

avec :

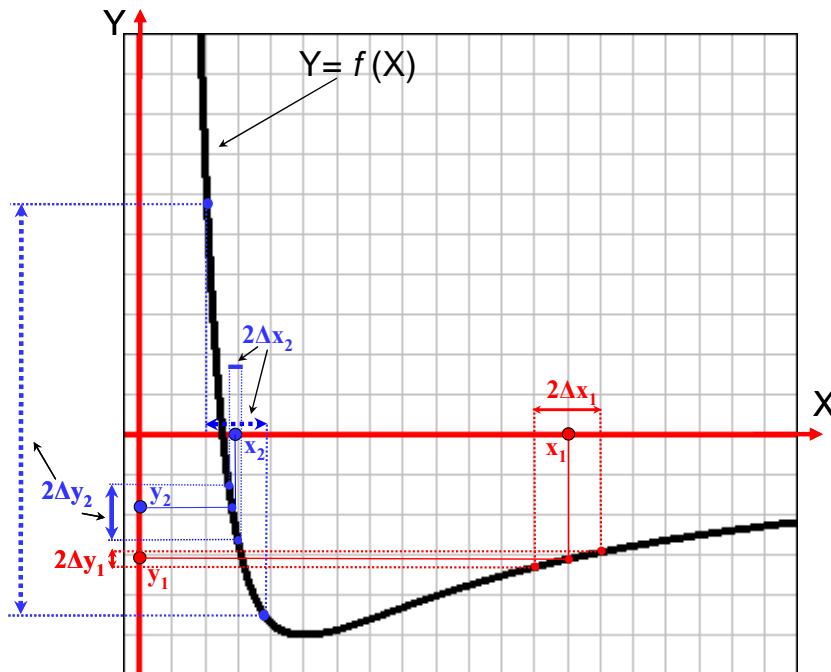
$$y = f(x)$$

$$\Delta y = |f'(x)| \Delta x$$

les unités de Y étant obtenues à partir de celles de X et de celles conférées à la fonction f .

³ Il faut faire appel ici au cours de mathématique et penser en particulier au développement de Taylor d'une fonction au voisinage de X : $f(X + dX) = f(X) + f'(X) dX + f''(X) (dX)^2 / 2! + \dots$, avec $dY = f(X + dX) - f(X)$. On voit que pour pouvoir écrire : $dY = f'(X) dX$, la quantité dX doit être suffisamment petite pour que tous les autres termes du développement de Taylor soient négligeables.

On notera que l'erreur relative $\Delta y/y = \Delta x |f'(x)|/f(x)$ dépend des variations de la fonction f , plus précisément des variations de sa dérivée logarithmique⁴. Une faible erreur relative sur X peut ainsi engendrer une erreur relative importante sur Y ; mais l'inverse peut survenir aussi, à savoir qu'une importante erreur relative sur X peut conduire à une erreur relative minime sur Y . Il en est de même pour les erreurs absolues comme le montre la figure ci-dessous :



Propagation de l'incertitude de mesure sur la valeur de $Y = f(X)$ après une mesure de X : les variations de la fonction f sont primordiales. 1. On mesure d'abord une valeur de X située sur la partie droite du graphique (x_1) avec une incertitude absolue Δx_1 : sur cette portion de la courbe, la valeur y_1 de Y est alors connue avec une bonne précision. 2. On mesure cette fois une valeur de X située sur la partie gauche (x_2). Si l'incertitude absolue de cette mesure est du même ordre de grandeur que la précédente, l'incertitude sur la valeur y_2 de Y est alors réhivitoire, au point que y_2 n'est même plus centrée sur l'intervalle $[f(x_2 + \Delta x_2) ; f(x_2 - \Delta x_2)]$: cela est dû au fait que, pour cette partie de la courbe, l'importance de Δx_2 est telle que les termes d'ordre supérieur à 1 dans le développement de Taylor de la fonction f au voisinage de x_2 (termes en $(\Delta x_2)^2$, etc..) ne sont pas négligeables. Sur cette partie de la courbe, il faut donc réduire l'incertitude absolue de la mesure de x_2 , quitte à changer d'appareil de mesure (ou de calibration), pour obtenir une incertitude plus raisonnable sur y_2 .

⁴ En effet : $[Log(f(x))]' = f'(x)/f(x)$

Interprétation statistique des incertitudes

Il a été reconnu au début du polycopié qu'une détermination précise de l'incertitude absolue d'une mesure n'est pas réalisable : l'estimation de cette valeur repose, après réflexion, sur un nombre bien trop important de conditions idéales à satisfaire que ni le constructeur de l'appareil de mesure, ni l'expérimentatrice ne peuvent affirmer qu'elles sont totalement remplies. Si bien que l'on finit toujours par déclarer une confiance raisonnable envers la mise en œuvre de la mesure effectuée et on place cette confiance dans l'estimation de Δg . L'interprétation statistique des incertitudes de mesure tente de donner un sens rationnel à cette « confiance raisonnable » au travers d'une analyse mathématique du probable.

Dans cette interprétation, on admet que la mesure est sujette à des fluctuations aléatoires. La grandeur physique réelle, G , lorsqu'elle est mesurée, prend des valeurs qui se répartissent continûment selon une certaine loi de probabilité. G prend le statut de *variable aléatoire* et le caractère continu des valeurs possibles que peut donner le résultat de la mesure conduit à la notion de *densité de probabilité*⁵, P , telle que $P(u)du$ soit la probabilité d'obtenir une valeur comprise dans l'intervalle $[u, u + du]$ lors d'une mesure de la grandeur physique G .

En général, on admet que la distribution des probabilités de mesure suit une *loi normale* représentée par une courbe en cloche appelée *gaussienne*⁶. Cette hypothèse est sous-tendue par le *théorème central limite* :

Soient n mesures de G qui ont fourni les résultats g_1, g_2, \dots, g_n , et considérons-les comme un échantillon de mesures aléatoires et indépendantes soumises à la même loi de probabilité (hypothèses qui ne sont pas toujours vérifiées dans la pratique). Soit μ l'espérance mathématique associée à cette loi de probabilité et σ l'écart type.

Soit maintenant g la valeur moyenne empirique des résultats de mesure :

$$g = \frac{1}{n} \sum_{i=1}^n g_i.$$

Alors, pour n suffisamment grand, g converge vers l'espérance mathématique μ (*loi des grands nombres*) en suivant une loi normale.

On a, par exemple, les encadrements suivants :

$$(g - \mu) \in \left[-\frac{\sigma}{\sqrt{n}}; +\frac{\sigma}{\sqrt{n}}\right] \quad \text{avec une probabilité de } 0.6826 = 68\%$$

$$(g - \mu) \in \left[-\frac{2\sigma}{\sqrt{n}}; +\frac{2\sigma}{\sqrt{n}}\right] \quad \text{avec une probabilité de } 0.9544 = 95\%$$

$$(g - \mu) \in \left[-\frac{3\sigma}{\sqrt{n}}; +\frac{3\sigma}{\sqrt{n}}\right] \quad \text{avec une probabilité de } 0.9974 = 99.7\%$$

Dans l'interprétation statistique des incertitudes de mesure, on se saisit du théorème central limite en passant outre le nombre de mesures effectuées : on assimile l'espérance mathématique μ de la loi normale à la valeur exacte g_e de la grandeur physique G , l'écart

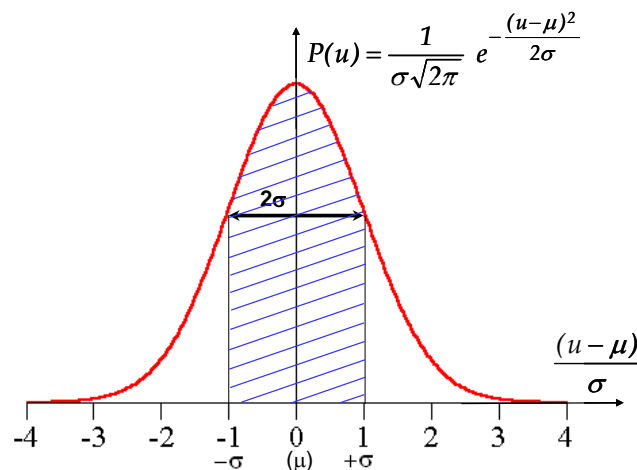
⁵ Voir le cours de mathématique sur les distributions statistiques.

⁶ Du nom de son découvreur : Karl Friedrich Gauss, 1777-1855, surnommé *le prince des mathématiciens*.

type σ à l'incertitude absolue Δg de la mesure, et la valeur mesurée g à la valeur moyenne empirique des résultats de mesure - quand bien même il n'y en aurait qu'une seule (ce que l'on avait déjà fait).

De la sorte, on est amené à considérer que le résultat g de la mesure de G a de grandes chances (68%) de se situer dans l'intervalle $[g_e - \Delta g; g_e + \Delta g]$, et on écrit : $g_e = g \pm \Delta g$. Ou encore - puisqu'alors $g = g_e \pm \Delta g$: que la valeur exacte de G a 68% de chances de se situer dans l'intervalle $[g - \Delta g; g + \Delta g]$.

Ainsi, du point de vue statistique, l'intervalle « rigide » $[g - \Delta g; g + \Delta g]$ est devenu un *intervalle de confiance* ; et toutes les questions sans réponse à propos de l'exactitude de la mesure, y compris le fait de n'avoir réalisé qu'une seule fois cette mesure⁷, sont reportées dans la valeur que l'on va donner à Δg considéré comme un écart type, et dans l'intervalle (à $\pm \sigma$, $\pm 2\sigma$, ...) que l'on va choisir pour exprimer le *taux de confiance* que l'on a dans le résultat de la mesure sous la forme d'une probabilité.



La loi normale $P(u)$ tracée en fonction de la variable réduite $(u - \mu)/\sigma$. Toute portion d'aire sous la courbe, prise entre deux valeurs limites données de u , exprime une probabilité : par exemple, l'aire hachurée = $\int_{\mu - \sigma}^{\mu + \sigma} P(u) du$ est la probabilité pour que l'on trouve, lors d'une mesure (i.e. lors d'un tirage), une valeur u comprise dans l'intervalle $[\mu - \sigma; \mu + \sigma]$; le calcul donne 0.68. (95.4% pour l'intervalle à $\pm 2\sigma$, 99.7% pour l'intervalle à $\pm 3\sigma$).

NB : l'aire totale sous la courbe est naturellement normalisée à 1 puisqu'elle représente la probabilité pour que l'on mesure une valeur u comprise dans \mathbb{R} .

On a donc : $\int_{-\infty}^{+\infty} P(u) du = 1$.

On a aussi, par définition : Espérance mathématique : $\mu \equiv \int_{-\infty}^{+\infty} u P(u) du$; Variance : $\sigma^2 \equiv \int_{-\infty}^{+\infty} (u - \mu)^2 P(u) du$

⁷ Une étude plus approfondie prenant en compte le nombre n de mesures effectuées amène à introduire un coefficient correcteur, t (coefficient de Student), qui dépend de n et du taux de confiance τ que l'on souhaite accorder aux résultats : l'intervalle de confiance s'écrit alors : $g_e \in [g - t(n, \tau)\sigma/\sqrt{n} ; g + t(n, \tau)\sigma/\sqrt{n}]$. Les valeurs de $t(n, \tau)$ sont tabulées. Par exemple, si l'on souhaite un taux de confiance $\tau = 68\%$ pour $n = 5$ mesures, on a : $t = 1.2$ (et t tend vers 1 lorsque n est grand). Pour un taux de confiance $\tau = 95\%$ et $n = 5$ mesures, on a cette fois $t = 2.8$ (et t tend vers 2 lorsque n est grand), etc...

Mesures indirectes à l'aide de plusieurs autres mesures

Il arrive que la grandeur physique G que l'on souhaite déterminer ne puisse être obtenue que de façon indirecte, grâce à la mesure d'un certain nombre d'autres grandeurs physiques X, Y, Z, \dots : $G = f(X, Y, Z, \dots)$. C'est le cas, par exemple, de la détermination d'une longueur d'onde lumineuse, λ , dans l'expérience des fentes de Young : $\lambda = IA/D$ où I est l'interfrange (distance entre deux maxima d'intensité lumineuse successifs sur l'écran), A est la distance séparant les deux fentes et D la distance entre l'écran et les deux fentes.

Calcul classique

On estime l'incertitude absolue commise sur la mesure indirecte de G en utilisant *le calcul différentiel pour les fonctions à plusieurs variables* : une variation élémentaire dG de G est la somme de toutes les variations élémentaires de cette grandeur engendrées par les variations élémentaires dX, dY, dZ, \dots de toutes les grandeurs physiques dont dépend G , chacune prise séparément :

$$dG = \frac{\partial f}{\partial X} dX + \frac{\partial f}{\partial Y} dY + \frac{\partial f}{\partial Z} dZ + \dots$$

où $\partial f / \partial X$ est la *dérivée partielle* de la fonction $f(X, Y, Z, \dots)$ par rapport à la variable X , les autres variables étant maintenues fixes lors du calcul de cette dérivée.

Par exemple : $d\lambda = \frac{A}{D} di + \frac{I}{D} dA - \frac{IA}{D^2} dD$.

Si les variables utilisées X, Y, Z, \dots sont *indépendantes* et que les mesures de chacune de ces grandeurs n'ont aucune corrélation entre elles, alors on aura :

$$g = f(x, y, z, \dots)$$
$$\Delta g = \left| \frac{\partial f}{\partial X}(x, y, z, \dots) \right| \Delta x + \left| \frac{\partial f}{\partial Y}(x, y, z, \dots) \right| \Delta y + \left| \frac{\partial f}{\partial Z}(x, y, z, \dots) \right| \Delta z + \dots = \Delta g_{class}$$

où x, y, z, \dots sont les valeurs mesurées des grandeurs X, Y, Z, \dots et $\Delta x, \Delta y, \Delta z, \dots$ leurs incertitudes absolues.

Par exemple :

$$\lambda = \frac{ia}{d} \quad \text{et} \quad \Delta \lambda_{class} = \frac{a}{d} \Delta i + \frac{i}{d} \Delta a + \frac{ia}{d^2} \Delta d$$

On notera le changement de signe dû au choix délibéré de prendre la somme des valeurs absolues et non pas des valeurs algébriques : on se conforme ici à la définition classique de Δg en tant que limite supérieure de l'erreur commise en prenant en compte l'éventualité du cas le plus défavorable où toutes les erreurs se sont accumulées sans se compenser (par exemple, les valeurs mesurées de I et A ont été surestimées tandis que la valeur de D a été sous-estimée).

Attention : si les variables X, Y, Z, \dots sont corrélées [exemple1 : $X = h(Y)$] ou si les mesures sont corrélées [exemple2 : l'appareil de mesure est construit de telle sorte qu'une incertitude Δx sur la mesure de X entraîne une incertitude $\Delta y = \alpha \Delta x$ sur la mesure de Y], alors il faut

intégrer ces corrélations dans le calcul différentiel avant de passer à la somme des valeurs absolues ; sinon, on surestime abusivement Δg :

$$\text{Exemple1 : } \Delta g = \left| \frac{\partial f}{\partial Y}(x, y, z, \dots) + h'(y) \frac{\partial f}{\partial X}(x, y, z, \dots) \right| \Delta y + \left| \frac{\partial f}{\partial Z}(x, y, z, \dots) \right| \Delta z + \dots$$

$$\text{Exemple2 : } \Delta g = \left| \frac{\partial f}{\partial X}(x, y, z, \dots) + \alpha \frac{\partial f}{\partial Y}(x, y, z, \dots) \right| \Delta x + \left| \frac{\partial f}{\partial Z}(x, y, z, \dots) \right| \Delta z + \dots$$

Calcul statistique

Dans l'interprétation statistique, on démontre que si les variables X, Y, Z, \dots sont indépendantes et leurs mesures décorrélatées, alors, pour un nombre suffisamment grand de mesures de ces grandeurs, la somme des produits croisés des incertitudes absolues (qui sont assimilées aux écarts types : $\Delta x \equiv \sigma_x, \Delta y \equiv \sigma_y, \dots$) s'annule en moyenne. En d'autres termes - et de façon mnémotechnique - il n'est besoin que d'additionner les *variances* (σ^2) associées aux différentes grandeurs mesurées, affectées des coefficients $\partial f / \partial X, \partial f / \partial Y, \dots$:

$$\sigma_g^2 = \left(\frac{\partial f}{\partial X}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial Y}\right)^2 \sigma_y^2 + \left(\frac{\partial f}{\partial Z}\right)^2 \sigma_z^2 + \dots$$

soit :

$$\Delta g = \sqrt{\left(\frac{\partial f}{\partial X}\right)^2 \Delta x^2 + \left(\frac{\partial f}{\partial Y}\right)^2 \Delta y^2 + \left(\frac{\partial f}{\partial Z}\right)^2 \Delta z^2 + \dots} = \Delta g_{stat}$$

$$g = f(x, y, z, \dots)$$

Par exemple :

$$\lambda = \frac{i a}{d} \quad \text{et} \quad \Delta \lambda_{stat} = \sqrt{\left(\frac{a}{d}\right)^2 \Delta i^2 + \left(\frac{i}{d}\right)^2 \Delta a^2 + \left(\frac{i a}{d^2}\right)^2 \Delta d^2} .$$

Comparée à la formule classique, la formule statistique fournit toujours une estimation de la valeur de Δg plus optimiste : $\Delta g_{stat} \leq \Delta g_{class}$. Ce résultat peut s'interpréter par le fait que, statistiquement, les cas défavorables envisagés dans l'analyse classique sont forts peu probables : ils ne sont donc pas à prendre en compte dans l'estimation de Δg mais plutôt en élargissant l'intervalle de confiance.

L'habitude semble être prise aujourd'hui d'utiliser Δg_{stat} . On aura bien conscience toutefois que les deux valeurs Δg_{stat} et Δg_{class} sont du même ordre de grandeur et que lorsqu'une seule mesure de X, Y, Z, \dots est réalisée, on se trouve bien loin de l'idéal statistique ; il faut alors apporter des corrections aux estimations chiffrées des taux confiance affichés (voir la note de la page - 8 -).

Vérification de lois physiques et/ou de modèles : notions

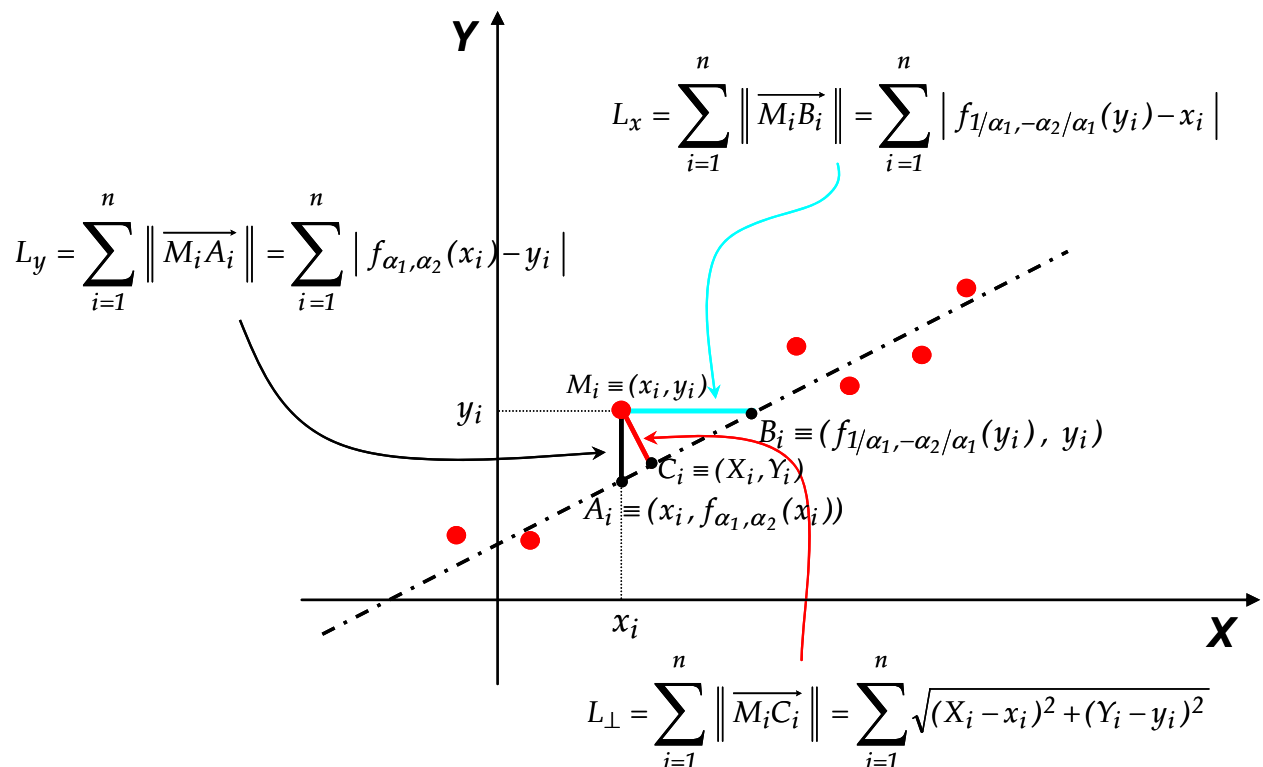
Introduction : méthode des moindres carrés

On peut mesurer différentes valeurs de deux grandeurs physiques X et Y dans le but de découvrir ou de vérifier si elles sont reliées entre elles par une relation fonctionnelle du type $Y = f(X)$. La question se pose alors de savoir comment influent les incertitudes expérimentales sur la recherche et sur la validation de la relation cherchée.

De façon courante, la fonction f est soit prévue par la théorie, soit inférée de l'expérience grâce à l'examen des résultats obtenus. Cette fonction fait partie en général d'une famille de fonctions qui se distinguent entre elles par les valeurs prises par k paramètres, $\alpha_1, \alpha_2, \dots, \alpha_k$, qui caractérisent chaque fonction de la famille : $f = f_{\alpha_1, \dots, \alpha_k}$

Supposons, par exemple, qu'après une série de n mesures $\{(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)\}$ il soit proposé une relation linéaire entre les grandeurs physiques X et Y : $Y = f_{\alpha_1, \alpha_2}(X) = \alpha_1 X + \alpha_2$. Il faut alors estimer les $k=2$ paramètres α_1 et α_2 qui permettent le meilleur ajustement possible des données expérimentales avec la famille de fonctions proposée.

Pour réaliser cet ajustement par le calcul, il faut se donner une mesure mathématique de l'écart entre les courbes proposées et les points expérimentaux ; la meilleure courbe étant alors celle qui minimise cette mesure. Le résultat dépend donc du choix de la mesure mathématique.



Pour fixer les idées, la figure ci-dessus montre trois choix de mesure possibles :

I. On peut vouloir minimiser par exemple la somme L_y des distances verticales qui séparent chaque point expérimental de la courbe proposée. Dans cette façon de voir, il est présupposé que chaque point expérimental M_i s'écarte du point « idéal » A_i de la courbe en

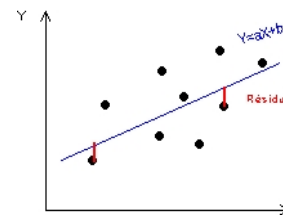
raison de la seule incertitude de mesure affectant y_i , la valeur de x_i étant supposée parfaitement connue.

2. On peut vouloir minimiser la somme L_x des distances horizontales qui séparent chaque point expérimental de la courbe proposée. Dans cette autre façon de voir, on présuppose cette fois que chaque point expérimental M_i s'écarte du point idéal B_i de la courbe à cause de la seule incertitude affectant x_i ; y_i étant parfaitement connue.

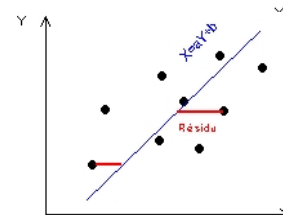
3. Si les incertitudes de mesure affectent tout autant les x_i que les y_i , alors on préférera minimiser la somme L_{\perp} des distances perpendiculaires entre chaque point M_i et la courbe proposée.

Dans chacun de ces trois cas, on est amené à calculer trois quantités différentes :

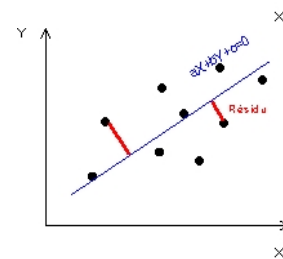
Cas n°1 :
$$L_y = \sum_{i=1}^n \left\| \overrightarrow{M_i A_i} \right\| = \sum_{i=1}^n \left| \alpha_1 x_i + \alpha_2 - y_i \right|$$



Cas n°2 :
$$L_x = \sum_{i=1}^n \left\| \overrightarrow{M_i B_i} \right\| = \sum_{i=1}^n \left| \frac{\alpha_1 x_i + \alpha_2 - y_i}{\alpha_1} \right|$$



Cas n°3 :
$$L_{\perp} = \sum_{i=1}^n \left\| \overrightarrow{M_i C_i} \right\| = \sum_{i=1}^n \frac{\left| \alpha_1 x_i + \alpha_2 - y_i \right|}{\sqrt{1 + \alpha_1^2}}$$



On minimise ensuite ces quantités en annulant simultanément leurs dérivées partielles par rapport à α_1 et α_2 ⁸ ; par exemple : $\partial L_y / \partial \alpha_1 = \partial L_y / \partial \alpha_2 = 0$.

Naturellement, on ne trouvera pas le même jeu optimal de paramètres $(\alpha_1; \alpha_2)$ selon que l'on a choisi la mesure L_x , L_y ou L_{\perp} . Il est donc important d'adapter la mesure la plus appropriée à la série de données envisagée.

En fait, la minimisation des trois quantités ci-dessus pose problème : la présence de la valeur absolue dans chacune de ces trois sommes empêche de fournir une solution analytique et

⁸ **Théorème** : les valeurs x_0, y_0, \dots qui rendent extrémale une fonction de plusieurs variables, $h(x, y, \dots)$, sont celles qui permettent d'annuler simultanément toutes les dérivées partielles de la fonction h par rapport à ces variables : $\partial h / \partial x \Big|_{x_0, y_0, \dots} = \partial h / \partial y \Big|_{x_0, y_0, \dots} = \dots = 0$. Dans notre cas, les paramètres $\alpha_1, \alpha_2, \dots, \alpha_k$ caractérisant la fonction f deviennent des variables pour les sommes L_x, L_y ou L_{\perp} : $L_{x,y,\perp} = L_{x,y,\perp}(\alpha_1, \alpha_2, \dots, \alpha_k)$.

seules des méthodes numériques permettent d'obtenir les valeurs optimales des paramètres $(\alpha_1; \alpha_2)$ qui minimisent ces sommes.

Pour palier à ce problème, on convient de calculer la somme des carrés des distances envisagées soit :

$$D_y^2 = \sum_{i=1}^n \left\| \overrightarrow{M_i A_i} \right\|^2 = \sum_{i=1}^n (\alpha_1 x_i + \alpha_2 - y_i)^2$$

$$D_x^2 = \sum_{i=1}^n \left\| \overrightarrow{M_i B_i} \right\|^2 = \sum_{i=1}^n \frac{(\alpha_1 x_i + \alpha_2 - y_i)^2}{\alpha_1^2}$$

$$D_{\perp}^2 = \sum_{i=1}^n \left\| \overrightarrow{M_i C_i} \right\|^2 = \sum_{i=1}^n \frac{(\alpha_1 x_i + \alpha_2 - y_i)^2}{1 + \alpha_1^2}$$

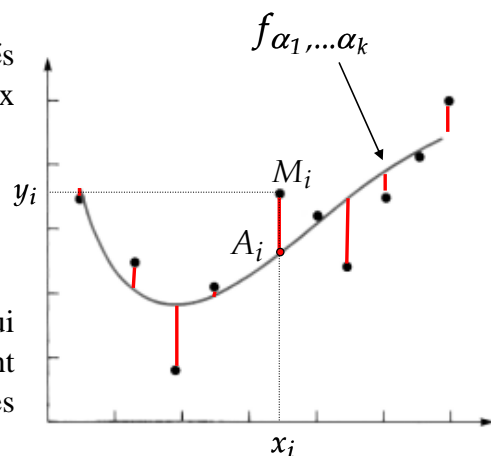
Cette fois, le calcul analytique des valeurs optimales $(\alpha_1; \alpha_2)$ est possible : c'est la méthode des moindres carrés.

De façon usuelle, la méthode consiste à estimer plutôt D_y^2 pour les raisons suivantes : premièrement, on montre que minimiser D_x^2 revient à minimiser un D_y^2 où l'on aurait interverti au préalable les x_i en les y_i moyennant une redéfinition des paramètres $\alpha_1, \alpha_2, \dots, \alpha_k$ (voir l'exemple plus bas). Deuxièmement, il s'avère que la minimisation de D_{\perp}^2 est difficile, car, hormis le cas simple abordé ci-dessous, on ne sait pas en général résoudre analytiquement les équations $\partial D_{\perp}^2 / \partial \alpha_1 = \partial D_{\perp}^2 / \partial \alpha_2 = \dots = \partial D_{\perp}^2 / \partial \alpha_k = 0$ parce qu'elles ne sont pas linéaires ; il faut donc recourir ici encore à des méthodes numériques.

En résumé, la méthode usuelle des moindres carrés consiste à évaluer la somme D_y^2 des écarts verticaux $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ élevés au carré :

$$D_y^2 = \sum_{i=1}^n |f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i|^2,$$

puis à rechercher les valeurs des k paramètres qui permettent de minimiser D_y^2 en annulant simultanément les dérivées partielles $\partial D_y^2 / \partial \alpha_1, \partial D_y^2 / \partial \alpha_2, \dots, \partial D_y^2 / \partial \alpha_k$.



Exemples d'ajustements linéaires par la méthode des moindres carrés

• Minimisation de D_y^2

Il suffit d'annuler simultanément les dérivées partielles de D_y^2 par rapport à α_1 et α_2 . On obtient ainsi un système de deux équations linéaires aux deux inconnues α_1 et α_2 :

$$\begin{cases} \alpha_1 \left(\sum_i x_i^2 \right) + \alpha_2 \left(\sum_i x_i \right) = \sum_i x_i y_i \\ \alpha_1 \left(\sum_i x_i \right) + n \alpha_2 = \sum_i y_i \end{cases}$$

La résolution de ce système fournit les valeurs uniques :

$$\alpha_1 = \frac{n \left(\sum x_i y_i \right) - \left(\sum x_i \right) \left(\sum y_i \right)}{n \left(\sum x_i^2 \right) - \left[\sum x_i \right]^2} \quad \text{et} \quad \alpha_2 = \frac{\left(\sum y_i \right) \left(\sum x_i^2 \right) - \left(\sum x_i \right) \left(\sum x_i y_i \right)}{n \left(\sum x_i^2 \right) - \left[\sum x_i \right]^2}$$

Ces formules sont en général directement implémentées dans les calculatrices ou les logiciels-tableurs.

• Minimisation de D_x^2

Le calcul des dérivées partielles de D_x^2 change à cause de la présence de α_1^2 au dénominateur. En fait, il y a avantage à considérer le problème en réécrivant D_x^2 de la façon suivante :

$$D_x^2 = \sum_{i=1}^n \left\| \overrightarrow{M_i B_i} \right\|^2 = \sum_{i=1}^n (y_i / \alpha_1 - \alpha_2 / \alpha_1 - x_i)^2$$

Posons alors : $x_i = y_i'$ et $y_i = x_i'$; et le problème revient à minimiser la quantité :

$$D_{y'}^2 = \sum_{i=1}^n (\alpha_1' x_i' + \alpha_2' - y_i')^2$$

avec : $\alpha_1' = 1/\alpha_1$ et $\alpha_2' = -\alpha_2/\alpha_1$.

Les valeurs optimales de α_1' et α_2' sont donc données par les formules du cas précédent où l'on a interverti les x_i en les y_i et réciproquement. On obtient ensuite : $\alpha_1 = 1/\alpha_1'$ et $\alpha_2 = -\alpha_2'/\alpha_1'$.

Il faut insister encore une fois sur le fait que, sauf exception, on n'obtiendra pas les mêmes valeurs de α_1 et α_2 que celles obtenues en minimisant D_y^2 .

Finalement, il est possible d'éviter le recours au calcul de D_x^2 en déclarant dès le départ de l'analyse que les y_i seront les données affectées par les incertitudes tandis que les x_i seront les données parfaites. De la sorte, on est toujours amené à minimiser un D_y^2 .

• *Minimisation de D_{\perp}^2*

La minimisation de D_{\perp}^2 ne se déduit pas de celle de D_y^2 puisque les x_i **et** les y_i sont maintenant affectées par les incertitudes de mesure.

On obtient une première équation en annulant la dérivée partielle de D_{\perp}^2 par rapport à α_2 :

$$\partial D_{\perp}^2 / \partial \alpha_2 = 2 \sum_i \frac{\alpha_1 x_i + \alpha_2 - y_i}{1 + \alpha_1^2} = 0 \quad \text{soit} \quad \alpha_2 = \frac{1}{n} \left(\sum_i y_i - \alpha_1 \sum_i x_i \right)$$

Le calcul de la dérivée partielle de D_{\perp}^2 par rapport à α_1 donne lieu à une seconde équation un peu plus compliquée. Finalement, après avoir injecté la valeur de α_2 dans cette équation, on trouve que α_1 doit être l'une des racines de l'équation du second degré :

$$\left(\sum_i x_i y_i \right) \alpha_1^2 + \left(\sum_i x_i^2 - \sum_i y_i^2 \right) \alpha_1 - \left(\sum_i x_i y_i \right) = 0$$

où l'on a posé : $x'_i = x_i - \frac{1}{n} \sum_i x_i$ et $y'_i = y_i - \frac{1}{n} \sum_i y_i$.

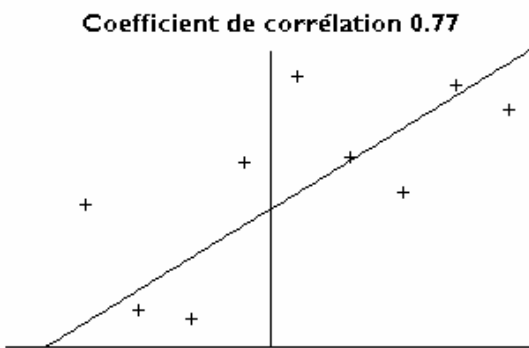
Attention : les calculettes et les logiciels-tableurs ne prennent pas toujours en compte ce cas de figure.

• *Qualité des ajustements*

La qualité de la *corrélation* linéaire entre les x_i et les y_i peut être quantifiée grâce au coefficient de corrélation linéaire dont la définition⁹ est :

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad \text{avec} \quad \bar{x} = \frac{1}{n} \sum x_i \quad \text{et} \quad \bar{y} = \frac{1}{n} \sum y_i$$

Les valeurs possibles de r sont comprises entre -1 et $+1$: une valeur de $+1$ (corrélation positive parfaite) est obtenue lorsque l'ensemble des couples de données (x_i, y_i) s'alignent parfaitement selon une pente positive ; la valeur -1 (corrélation négative parfaite) est obtenue pour un alignement parfait avec une pente négative (les y_i décroissent cependant que les x_i croissent).



La valeur $r=0$ indique que l'ensemble des couples (x_i, y_i) montre une absence totale de corrélation linéaire.

On ne doit cependant pas accorder un trop grand crédit au coefficient r pour juger de la relation de linéarité à laquelle obéirait l'ensemble des données comme le montre l'exemple de la figure ci-contre. *Ce coefficient ne représente qu'une tendance moyenne entre les données et le simple examen*

⁹ L'origine de cette définition est due au fait que les deux séries de valeurs $\vec{X} \equiv (x_1, x_2, \dots, x_n)$ et $\vec{Y} \equiv (y_1, y_2, \dots, y_n)$ peuvent être vues comme deux vecteurs d'un espace vectoriel de dimension n . Le coefficient r représente alors le cosinus de l'angle entre les deux vecteurs n -dimensionnels $(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$ et $(y_1 - \bar{y}, y_2 - \bar{y}, \dots, y_n - \bar{y})$: si $r = +1$, l'angle entre ces deux vecteurs est nul et ils sont colinéaires ; si $r = 0$ les deux vecteurs sont orthogonaux ; si $r = -1$ ils sont colinéaires et de sens opposé.

visuel d'un graphique en dit plus long que la valeur de ce coefficient. Enfin, une erreur courante est de croire qu'un coefficient de corrélation élevé induit forcément une relation de causalité entre les deux phénomènes mesurés. Ça n'est pas le cas : les deux phénomènes (les x_i et les y_i) peuvent tout aussi bien être corrélés à un même phénomène-source, c'est-à-dire à une troisième variable non mesurée, et dont dépendent les deux autres ¹⁰.

La qualité et la validité de l'ajustement peuvent aussi être estimées par différents moyens d'analyse statistique des *résidus*, qui sont les écarts résiduels $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ obtenus pour chaque couple de points (x_i, y_i) , après avoir remplacé les α_i par leurs valeurs optimales. L'analyse de ces résidus dépend fortement de la nature des données (x_i, y_i) .

La méthode des moindres carrés reste insatisfaisante au sens où elle se contente de fournir, parmi toutes les courbes proposées, celle qui passe au plus près de l'ensemble des points $\{(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)\}$ au vu de la mesure considérée ; mais elle ne prend pas en compte de façon quantitative les incertitudes de mesure et on ne sait pas dire si, pour chaque couple (x_i, y_i) , l'écart résiduel $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ est trop grand, trop petit, raisonnable ou aberrant. Pour donner un sens à ces qualificatifs, il faut pouvoir comparer cette différence avec les incertitudes expérimentales de chacun des couples de valeurs mesurées (x_i, y_i) : il faut donc rapporter l'écart $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ aux incertitudes de mesure Δy_i et/ou Δx_i .

La *méthode du khi-deux* qui va suivre repose sur l'interprétation statistique des erreurs de mesures en assimilant les incertitudes absolues aux écarts types : $\Delta x_i \equiv \sigma_{x_i}$, $\Delta y_i \equiv \sigma_{y_i}$, L'écart $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ apparaît alors comme une distribution statistique gaussienne centrée (i.e. de valeur moyenne nulle).

Méthode du khi-deux

Souvent, les x_i sont connus avec une précision telle que les σ_{x_i} associés sont négligeables. On rapporte donc la différence $f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i$ au seul écart-type σ_{y_i} en estimant, pour chaque couple (x_i, y_i) , la quantité : $[f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i] / \sigma_{y_i}$. On calcule ensuite la somme des carrés de ces quantités :

$$\chi_y^2 = \sum_{i=1}^n \left[\frac{f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i}{\sigma_{y_i}} \right]^2,$$

que l'on minimise par rapport aux α_j comme précédemment : c'est la *méthode du khi-deux*.

Ici, la mesure mathématique se trouve affectée d'un *poids* ($[1/\sigma_{y_i}]^2$) d'autant plus important que la mesure est précise ; chaque terme de la somme se voyant pondéré de la sorte, la minimisation qui va s'ensuivre sera naturellement plus sensible aux termes prépondérants, c'est-à-dire aux valeurs mesurées les plus précises.

¹⁰ par exemple : à supposer que l'on constate une relation linéaire entre le nombre de personnes à Nancy qui portent des lunettes de soleil la journée du 15 août et le nombre de personnes qui ont été à la piscine ce même jour, que penser de cette corrélation ?

Lorsque les incertitudes absolues (les écarts types σ_{x_i}) sur les x_i sont considérées comme non négligeables, on adopte la *méthode du khi-deux généralisée* (ou *perpendiculaire*) qui consiste à additionner les variances associées à chaque couple (x_i, y_i) - compte tenu de la relation supposée : $Y = f_{\alpha_1, \dots, \alpha_k}(X)$ - et à pondérer chaque terme de la somme par le résultat ainsi obtenu, soit :

$$\chi_{\perp}^2 = \sum_{i=1}^n \frac{[f_{\alpha_1, \dots, \alpha_k}(x_i) - y_i]^2}{\sigma_{y_i}^2 + [\sigma_{x_i} \frac{df_{\alpha_1, \dots, \alpha_k}}{dX}(x_i)]^2}$$

L'exemple des ajustements linéaires qui va suivre aidera sans doute à mieux saisir de façon intuitive pourquoi le dénominateur prend cette forme.

On minimise ensuite le χ_{\perp}^2 par rapport aux α_j comme précédemment. Il faut noter la présence de la dérivée de la fonction f au dénominateur qui complique sérieusement la recherche des k paramètres $\alpha_1, \alpha_2, \dots, \alpha_k$ qui minimisent ce χ_{\perp}^2 .

Exemples d'ajustements linéaires par la méthode du khi-deux

- Minimisation de χ_y^2

Ici, la mesure utilisée pour l'ajustement linéaire vertical est celle du « khi-deux-y » :

$$\chi_y^2 = \sum_{i=1}^n \left[\frac{\alpha_1 x_i + \alpha_2 - y_i}{\sigma_{y_i}} \right]^2.$$

Sa minimisation par rapport à α_1 et α_2 donne lieu à un système d'équations linéaires aux deux inconnues α_1 et α_2 :

$$\begin{cases} \alpha_1 \left(\sum_i x_i^2 / \sigma_{y_i}^2 \right) + \alpha_2 \left(\sum_i x_i / \sigma_{y_i}^2 \right) = \sum_i x_i y_i / \sigma_{y_i}^2 \\ \alpha_1 \left(\sum_i x_i / \sigma_{y_i}^2 \right) + \alpha_2 \left(\sum_i 1 / \sigma_{y_i}^2 \right) = \sum_i y_i / \sigma_{y_i}^2 \end{cases}$$

On obtient ainsi la solution unique :

$$\alpha_1 = \frac{\left(\sum 1 / \sigma_{y_i}^2 \right) \left(\sum x_i y_i / \sigma_{y_i}^2 \right) - \left(\sum x_i / \sigma_{y_i}^2 \right) \left(\sum y_i / \sigma_{y_i}^2 \right)}{\left(\sum 1 / \sigma_{y_i}^2 \right) \left(\sum x_i^2 / \sigma_{y_i}^2 \right) - \left[\sum x_i / \sigma_{y_i}^2 \right]^2}$$

et

$$\alpha_2 = \frac{\left(\sum y_i / \sigma_{y_i}^2 \right) \left(\sum x_i^2 / \sigma_{y_i}^2 \right) - \left(\sum x_i / \sigma_{y_i}^2 \right) \left(\sum x_i y_i / \sigma_{y_i}^2 \right)}{\left(\sum 1 / \sigma_{y_i}^2 \right) \left(\sum x_i^2 / \sigma_{y_i}^2 \right) - \left[\sum x_i / \sigma_{y_i}^2 \right]^2}$$

Il faut constater que lorsque tous les σ_{y_i} sont égaux à une seule et même valeur, σ_y , alors les solutions sont celles obtenues précédemment par la méthode des moindres carrés : la méthode

des moindres carrés revient donc à admettre que tous les y_i sont déterminés avec la même incertitude absolue σ_y et, en conséquence, à donner un même poids statistique à toutes les mesures.

La méthode du khi-deux permet également d'évaluer les incertitudes absolues (leurs écarts-types) sur les deux valeurs α_1 et α_2 fournies par le calcul précédent. En différenciant les deux formules ci-dessus par rapport aux y_i , on démontre les formules suivantes :

$$\sigma_{\alpha_1}^2 = \frac{\sum x_i^2 / \sigma_{y_i}^2}{\left(\sum 1 / \sigma_{y_i}^2\right) \left(\sum x_i^2 / \sigma_{y_i}^2\right) - \left[\sum x_i / \sigma_{y_i}^2\right]^2} \quad \text{et} \quad \sigma_{\alpha_2}^2 = \frac{\sum 1 / \sigma_{y_i}^2}{\left(\sum 1 / \sigma_{y_i}^2\right) \left(\sum x_i^2 / \sigma_{y_i}^2\right) - \left[\sum x_i / \sigma_{y_i}^2\right]^2}$$

Lorsque $\sigma_{y_i} = \sigma_y \forall i$, ces deux formules permettent d'attribuer des incertitudes absolues aux deux coefficients α_1 et α_2 déterminés par la méthode des moindres carrés. On aura conscience toutefois que l'ajustement obtenu à l'aide de la méthode des moindres carrés est indépendant de la valeur attribuée à σ_y

• *Minimisation de χ_x^2*

On ne sera pas étonné de la définition du « khi-deux-x » lorsque les incertitudes absolues affectent seulement les données x_i :

$$\chi_x^2 = \sum_{i=1}^n \left[\frac{\alpha_1 x_i + \alpha_2 - y_i}{\alpha_1 \sigma_{x_i}} \right]^2 .$$

C'est qu'en effet, une incertitude absolue σ_{x_i} sur la valeur de x_i entraîne une incertitude $\alpha_1 \sigma_{x_i}$ sur la valeur de $f_{\alpha_1, \alpha_2}(x_i) = \alpha_1 x_i + \alpha_2$.

Comme pour le traitement de D_x^2 , il y a avantage à considérer le problème en réécrivant χ_x^2 de la façon suivante :

$$\chi_x^2 = \sum_{i=1}^n \left[\frac{y_i / \alpha_1 - \alpha_2 / \alpha_1 - x_i}{\sigma_{x_i}} \right]^2 ;$$

puis à poser : $x_i = y'_i$, $\sigma_{x_i} = \sigma_{y'_i}$ et $y_i = x'_i$, de sorte que le problème revient à minimiser la quantité :

$$\chi_{y'}^2 = \sum_{i=1}^n \left[\frac{\alpha'_1 x'_i + \alpha'_2 - y'_i}{\sigma_{y'_i}} \right]^2$$

avec : $\alpha'_1 = 1/\alpha_1$ et $\alpha'_2 = -\alpha_2/\alpha_1$.

Les valeurs optimales de α'_1 et α'_2 ainsi que leurs incertitudes absolues sont donc données par les formules précédentes où l'on a interverti les x_i en les y_i et réciproquement. On obtient ensuite : $\alpha_1 = 1/\alpha'_1$ et $\alpha_2 = -\alpha'_2/\alpha'_1$.

- *Minimisation de χ^2_{\perp}*

Si les incertitudes affectent à la fois les x_i et les y_i sans que l'on puisse négliger leurs influences respectives, il faut alors prendre la mesure du khi-deux généralisé qui s'écrit ici :

$$\chi^2_{\perp} = \sum_{i=1}^n \frac{[\alpha_1 x_i + \alpha_2 - y_i]^2}{\sigma_{y_i}^2 + [\sigma_{x_i} \alpha_1]^2}$$

où cette fois le poids statistique associé à chaque couple (x_i, y_i) est représenté par l'inverse de la somme des variances associées à chacune des incertitudes. Les équations de minimisation du χ^2_{\perp} par rapport aux α_i s'écrivent maintenant :

$$\frac{\partial \chi^2}{\partial \alpha_1} = \sum_i \frac{\partial}{\partial \alpha_1} \left[\frac{[\alpha_1 x_i + \alpha_2 - y_i]^2}{\sigma_{y_i}^2 + [\sigma_{x_i} \alpha_1]^2} \right] = 0 \quad \text{et} \quad \frac{\partial \chi^2}{\partial \alpha_2} = \sum_i \frac{2[\alpha_1 x_i + \alpha_2 - y_i]}{\sigma_{y_i}^2 + [\sigma_{x_i} \alpha_1]^2} = 0$$

De graves ennuis surviennent pour résoudre ces équations : la présence de α_1 au dénominateur empêche sa factorisation et ne laisse que l'espoir de résoudre ces équations par un traitement numérique.

- *Qualité des ajustements*

Comme pour la méthode des moindres carrés, on a l'habitude d'estimer la qualité de l'ajustement des données par la fonction $f_{\alpha_1, \dots, \alpha_k}$ en analysant les résidus de l'ajustement avec différents outils statistiques qui dépendent la nature des données expérimentales.

La coutume, fondée sur l'analyse statistique, est de considérer l'ajustement valide si $\chi^2 \cong n - k$ avec un écart-type admissible $\Delta \chi^2 = \sqrt{2(n - k)}$.

On notera que si $n = k$, alors le χ^2 est nul : c'est qu'en effet on peut toujours trouver k paramètres $\alpha_1, \alpha_2, \dots, \alpha_k$ tels que la fonction $f_{\alpha_1, \dots, \alpha_k}$ passe automatiquement par les $n = k$ points expérimentaux (si $k = 2$: par 2 points passe toujours une droite, on ne peut donc pas discuter de la qualité de l'ajustement dans ce cas). Afin que l'ajustement ait un sens il faut donc que le nombre de points expérimentaux soit plus grand que le nombre de paramètres à ajuster. On appelle la quantité $n - k$ le *nombre de degrés de liberté* du problème.