ANNEXE

Meilleur estimateur de $\overline{x}$

Montrons que le meilleur estimateur de $\overline{x}$ lorsque la densité de probabilité est une gaussienne est $x^* = \frac{1}{N}\sum_{i=1}^{N} x_i$.

On a réalisé N mesures de X donnant les valeurs x_i .
La probabilité d'obtenir la valeur x_i (à dx près) effectivement obtenue lors de la $i^{\text{ème}}$ mesure est proportionnelle à $e^{-\frac{(x_i-\overline{x})^2}{2\sigma^2}}$.
La probabilité P d'obtenir l'ensemble des valeurs x_i (événements indépendants) est donc proportionnelle à $\prod_{i=1}^{N} e^{-\frac{(x_i-\overline{x})^2}{2\sigma^2}} = \exp\left[-\frac{\sum_{i=1}^{N}(x_i-\overline{x})^2}{2\sigma^2}\right]$

On applique alors le **principe du maximum de vraisemblance** : la valeur cherchée $\overline{x}$ (inconnue) est estimée en cherchant, parmi toutes les valeurs possibles de $\overline{x}$, celle x^* qui rend maximale la probabilité P d'obtenir l'ensemble des valeurs x_i .
En effet, la probabilité d'avoir cet ensemble de mesures est un produit de lois gaussiennes : il y a beaucoup de combinaisons possibles, donc beaucoup de résultats expérimentaux possibles, mais tous correspondent pratiquement à la probabilité maximale, et ce d'autant plus que N est grand.

Ceci entraîne $\frac{\mathrm{d}}{\mathrm{d}\overline{x}}\left[\sum_{i=1}^{N}(x_i-\overline{x})^2\right] = -2\sum_{i=1}^{N}(x_i-\overline{x}) = 0$ soit $\boxed{\overline{x} = x^* = \frac{1}{N}\sum_{i=1}^{N} x_i}$

Justification de $\sigma_{x^*} = \frac{\sigma}{\sqrt{N}}$

Réalisons une série de N mesures indépendantes de la grandeur X : elle fournit des résultats x_i avec $i \in [1,N]$. On en déduit la valeur moyenne de la série $x^* = \frac{\sum_{i=1}^{N} x_i}{N}$.
Les mesures x_i sont distribuées selon une loi normale dont l'écart-type est σ.

La formule de propagation des incertitudes fournit :

$$\boxed{\sigma_{x^*} = \frac{\sqrt{\sum_{i=1}^{N}\sigma^2}}{N} = \frac{\sigma\sqrt{N}}{N} = \frac{\sigma}{\sqrt{N}}}.$$

Plus on réalise de mesures, plus l'incertitude sur la valeur moyenne est faible.

Justification de $\sigma^* = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - x^*)^2}$ (ou : pourquoi $N-1$ et pas N au dénominateur ?)

On peut recommencer un très grand nombre de fois (en théorie une infinité) des séries de N mesures $x_{i,k}$: pour la k ème série, on obtient $x_k^* = \frac{\sum_{i=1}^{N} x_{i,k}}{N}$.

Soit $A_k = \sum_{i=1}^{N}(x_{i,k} - x_k^*)^2$. On veut donc montrer que $A = \overline{A_k} = (N-1)\sigma^2$

Or $A_k = \sum_{i=1}^{N} x_{i,k}{}^2 - 2\sum_{i=1}^{N} x_{i,k}x_k{}^* + \sum_{i=1}^{N} x_k{}^{*2} = \sum_{i=1}^{N} x_{i,k}{}^2 - 2Nx_k{}^{*2} + Nx_k{}^{*2}$ soit :

$$A_k = \sum_{i=1}^{N} x_{i,k}{}^2 - Nx_k{}^{*2} = \sum_{i=1}^{N} x_{i,k}{}^2 - \frac{1}{N}\left[\sum_{i=1}^{N} x_{i,k}\right]^2 = \sum_{i=1}^{N} x_{i,k}{}^2 - \frac{1}{N}\left[\sum_{i=1}^{N} x_{i,k}\right]\left[\sum_{j=1}^{N} x_{j,k}\right]$$

$$A_k = \sum_{i=1}^{N} x_{i,k}{}^2 - \frac{1}{N}\left[\sum_{i=1}^{N} x_{i,k}{}^2\right] - \frac{1}{N}\left[\sum_{\substack{i,j\\ i\neq j}}^{N} x_{i,k}x_{j,k}\right]$$

D'où $\overline{A_k} = \frac{N-1}{N}\sum_{i=1}^{N}\overline{x_{i,k}{}^2} - \frac{1}{N}\left[\sum_{\substack{i,j\\ i\neq j}}^{N}\overline{x_{i,k}x_{j,k}}\right]$

Or $\sigma^2 = \overline{(x_i - \overline{x})^2} = \overline{x_i{}^2} - \overline{x}^2 \Rightarrow \overline{x_{i,k}{}^2} = \overline{x_i{}^2} = \sigma^2 + \overline{x}^2$

D'autre part, les mesures étant indépendantes,

$\overline{(x_i - \overline{x})(x_j - \overline{x})} = 0 \Leftrightarrow \overline{(x_i x_j - x_i\overline{x} - x_j\overline{x} + \overline{x}^2)} = 0$

soit $\overline{x_i x_j} - \overline{x}^2 - \overline{x}^2 + \overline{x}^2 = \overline{x_i x_j} - \overline{x}^2 = 0$ d'où $\overline{x_{i,k}x_{j,k}} = \overline{x_i x_j} = \overline{x}^2$

On a finalement bien $A = \frac{N-1}{N}N(\sigma^2 + \overline{x}^2) - \frac{1}{N}N(N-1)\overline{x}^2$, soit :

$$\boxed{A = (N-1)(\sigma^2 + \overline{x}^2) - (N-1)\overline{x}^2 = (N-1)\sigma^2}$$

Propagation des incertitudes

La mesure d'une grandeur X fournit un résultat $\overline{x}$ avec une incertitude Δx.

Supposons que la grandeur X ne soit pas directement accessible par la mesure mais liée à des mesures **indépendantes** de **deux grandeurs elles-mêmes indépendantes** X_1 et X_2 dont les valeurs mesurées sont $\overline{x_1}$ (avec un écart-type σ_1, ou une incertitude élargie Δx_1) et $\overline{x_2}$ (avec un écart-type σ_2, ou une incertitude élargie Δx_2) : $X = f(X_1, X_2)$.

Si les incertitudes sont supposées petites, on peut effectuer un développement limité à l'ordre 1 de la fonction f autour du point $(\overline{x_1}, \overline{x_2})$:

$$x = f(x_1, x_2) = f(\overline{x_1}, \overline{x_2}) + (x_1 - \overline{x_1})\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2}) + (x_2 - \overline{x_2})\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})$$

où $\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2})$ et $\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})$ sont les dérivées partielles de X par rapport à X_1 et X_2, calculées au point $(\overline{x_1}, \overline{x_2})$.

Prenons la valeur moyenne de cette expression :

$\overline{x} = f(\overline{x_1}, \overline{x_2}) + \overline{(x_1 - \overline{x_1})}\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2}) + \overline{(x_2 - \overline{x_2})}\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2}) = f(\overline{x_1}, \overline{x_2})$, puisque : $\overline{(x_1 - \overline{x_1})} = \overline{x_1} - \overline{x_1} = 0$ et $\overline{(x_2 - \overline{x_2})} = \overline{x_2} - \overline{x_2} = 0$, les valeurs mesurées étant par définition les valeurs moyennes.

On a donc :

$$\boxed{\overline{x} = f(\overline{x_1}, \overline{x_2})}$$

Il reste à déterminer l'incertitude sur *x* (due aux incertitudes sur x_1 et x_2).

Par définition de l'écart type, on a $\sigma^2 = \overline{(x - \overline{x})^2} = \overline{\left[x - f(\overline{x_1}, \overline{x_2})\right]^2}$ d'où :

$$\sigma^2 = \overline{(x_1 - \overline{x_1})^2}\left[\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2})\right]^2 + \overline{(x_2 - \overline{x_2})^2}\left[\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})\right]^2 + 2\overline{(x_1 - \overline{x_1})(x_2 - \overline{x_2})}\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2})\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})$$

Les variables X_1 et X_2 étant indépendantes, on a $\overline{(x_1 - \overline{x_1})(x_2 - \overline{x_2})} = 0$ (corrélation nulle).

Comme par définition $\overline{(x_1 - \overline{x_1})^2} = \sigma_1^2$ et $\overline{(x_2 - \overline{x_2})^2} = \sigma_2^2$, on a :

$$\sigma = \sqrt{\sigma_1^2\left[\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2})\right]^2 + \sigma_2^2\left[\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})\right]^2}$$

L'incertitude élargie étant proportionnelle à l'écart-type, on obtient la **relation fondamentale permettant de calculer l'incertitude sur *X*** :

$$\boxed{\Delta x = \sqrt{(\Delta x_1)^2\left[\frac{\partial f}{\partial X_1}(\overline{x_1}, \overline{x_2})\right]^2 + (\Delta x_2)^2\left[\frac{\partial f}{\partial X_2}(\overline{x_1}, \overline{x_2})\right]^2}}$$

Écart-type d'un densité de probabilité rectangulaire.

La densité de probabilité est supposée uniforme, égale à a, entre x_1 et x_2, et nulle hors de cet intervalle.

On a donc $\int_{x_1}^{x_2} a\mathrm{d}x = 1 \Rightarrow a = \frac{1}{x_2 - x_1}$

On en déduit la valeur moyenne $\overline{x} = \int_{x_1}^{x_2} ax\mathrm{d}x = \frac{a}{2}(x_2{}^2 - x_1{}^2) = \frac{x_1 + x_2}{2}$ et l'écart-type :

$$\sigma^2 = \int_{x_1}^{x_2} a(x - \overline{x})^2\mathrm{d}x = \int_{x_1}^{x_2} ax^2\mathrm{d}x - \overline{x}^2 = \frac{a}{3}(x_2{}^3 - x_1{}^3) - \overline{x}^2$$

$$= \frac{x_1{}^2 + x_1x_2 + x_2{}^2}{3} - \frac{x_1{}^2 + 2x_1x_2 + x_2{}^2}{4} = \frac{x_1{}^2 - 2x_1x_2 + x_2{}^2}{12}$$

$$= \frac{(x_2 - x_1)^2}{12}$$

D'où $\boxed{\sigma = \frac{x_2 - x_1}{\sqrt{12}}}$

Bibliographie sommaire :

TAYLOR John, *Incertitudes et analyse des erreurs dans les mesures physiques*, Dunod, 2000, 318 p.

BELORIZKY Elie, *Probabilités et statistiques dans les sciences expérimentales*, Nathan Université, collection 128, 1998, 130 p.

VENTZEL Hélène, *Théorie des probabilités*, Editions Mir, 1982, 572 p.

MOREAU René, *Mesures, erreurs et incertitudes en Physique-Chimie*, Université d'Angers [en ligne]. Disponible (et également sur d'autres sites)
http://ead.univ-angers.fr/~capespc/physique/generalites/mesureserreursincertitudes_moreau2.pdf

Groupe des Sciences physiques et chimiques de l'IGEN, *Nombres, mesures et incertitudes en sciences physiques et chimiques*, EDUSCOL [en ligne]. Disponible
http://media.eduscol.education.fr/file/PC/66/3/Ressources_PC_nombres_mesures_incertitudes_144663.pdf