

Eléments de métrologie :
incertitudes et analyse des résultats de mesure

David Boilley¹
Université de Caen Basse-Normandie et GANIL

Ecole doctorale SIMEM
Année 2010-2011

¹GANIL, BP 55027, F-14076 Caen cedex 05,
<http://www.physique.unicaen.fr/~boilley>

boilley@ganil.fr,

tél : 02 31 45 47 81,

Table des matières

1	Présentation des résultats	4
1.1	Unités	4
1.1.1	Le système international	4
1.1.2	Symboles	4
1.2	Ecriture des nombres	4
1.2.1	Nombres	5
1.2.2	Règles d'écriture et d'arrondi	5
1.3	Vocabulaire	6
1.4	Le concept d'incertitude	6
1.4.1	Historique	6
1.4.2	Les processus de mesure et les processus d'essai	7
1.4.3	Facteurs d'influence	7
2	Estimation de l'incertitude suivant la méthode du GUM	9
2.1	Quand le résultat de mesure est une variable aléatoire	9
2.2	Détermination expérimentale de la variance : évaluation de type A de l'incertitude	11
2.2.1	Estimateur de la variance	11
2.2.2	Mise en œuvre pratique	13
2.3	Evaluation de type B de l'incertitude	14
2.4	Propagation de l'incertitude	14
2.4.1	Méthode	14
2.4.2	Exemples avec des grandeurs indépendantes	15
2.4.3	Prise en compte des corrélations éventuelles	16
2.5	Niveaux de confiance et degrés de liberté	19
2.5.1	Problématique	19
2.5.2	Table ou loi de Student	20
2.5.3	Nombre de degrés de liberté	20
2.6	Recommandations pour la présentation des résultats	22
2.7	Résumé et conclusion	22
3	Calcul de la propagation de l'incertitude en utilisant des méthodes Monte Carlo	24
3.1	Démarche	24
3.1.1	Formulation du problème	24
3.1.2	Les étapes de l'évaluation	25
3.2	Choix et construction des distributions d'entrée	26
3.2.1	Distribution rectangulaire	26
3.2.2	Distribution trapézoïdale	26
3.2.3	Distribution gaussienne simple	26
3.2.4	Distribution gaussienne à plusieurs variables	27
3.2.5	Générateur de nombres pseudo-aléatoires	27
3.3	Mise en œuvre pratique	28
3.3.1	Tolérance numérique	28
3.3.2	Stabilité numérique	28

3.3.3	Algorithme	28
3.4	Validation du GUM	29
4	Traitement statistique des données	30
4.1	Préambule : quelques notions sur les probabilités	30
4.2	Moyenne pondérée	31
4.3	Méthode des moindres carrés	32
4.3.1	Régression linéaire	32
4.3.2	Avec une incertitude sur les x_i et y_i	33
4.3.3	Autres fonctions	33
4.3.4	Cas d'une gaussienne	33
4.4	Le test du χ^2	34
4.4.1	Approche intuitive	34
4.4.2	Approche quantitative	34
5	Seuil de décision, limite de détection et intervalle de confiance	36
5.1	Introduction	36
5.2	Avec une distribution de Poisson	37
5.2.1	Distribution de Poisson	37
5.2.2	Seuil de décision	38
5.2.3	Limite de détection	40
5.2.4	Calcul de l'activité	42
5.2.5	Extrémités de l'intervalle de confiance et meilleure estimation	43
5.3	Exemples d'optimisation du seuil de décision	44
5.3.1	Influence des durées de comptage	44
5.3.2	Extraction du bruit de fond d'un spectre	45
5.4	Cas général	45
5.5	Limite de décision	45
6	Méthodes bayésiennes de calcul de l'incertitude	47
6.1	Un peu de théorie des probabilités	47
6.1.1	Un peu d'algèbre	47
6.1.2	Probabilité conditionnelle	48
6.1.3	Théorème de Bayes	49
6.1.4	Inférence bayésienne	49
6.2	Maximisation de l'entropie de l'information	50
6.3	Application aux incertitudes	52
6.3.1	Retour sur les seuils de décision et limites de détection	52
6.3.2	Evaluation de type A de l'incertitude	53
A	Historique de la normalisation du calcul d'incertitude	54

Introduction

Les résultats des mesures que vous allez faire doivent pouvoir être utilisés par tous, tout comme vous devez pouvoir utiliser ceux obtenus par d'autres. C'est le cas pour l'étalonnage par exemple. Au-delà d'un système d'unité commun à tous, il faut définir une convention commune dans l'expression des résultats. Cela commence par le vocabulaire, l'expression mathématique et bien entendu le calcul d'incertitude, qui va être le gros morceau de ce cours. La logique est que l'information soit claire, complète (voire redondante) et exploitable. Pour cela, il existe tout un corpus de normes à respecter que vous devez connaître.

L'analyse de l'incertitude et des erreurs liées à un mesurage est fondamentale en métrologie. C'est en effet nécessaire pour progresser. L'estimation de l'incertitude d'un mesurage est souvent plus dur que le mesurage lui-même. Elle nécessite de bien connaître le processus de mesure dans ses moindres détails et quelques notions de statistique. Ce cours a pour ambition de vous apporter ces notions de statistique et de faciliter la lecture des normes et des ouvrages plus spécialisés. Il ne vous dispense pas de leur lecture.

Il s'agit de la première version de ces notes de cours. Elle contient sûrement de nombreux bugs. Si vous avez des remarques, critiques, commentaires... ils sont les bienvenus.

Chapitre 1

Présentation des résultats

Ce premier chapitre est basé sur les normes suivantes :

- Norme française NF X 02-003 : **Principe de l'écriture des nombres, des grandeurs, des unités et des symboles**, décembre 1995.
- Norme française NF X 02-006 : **Le système international d'unités**, août 1994.
- Norme française NF X 07-001 : **Vocabulaire international des termes fondamentaux et généraux de métrologie**, décembre 1994.

Ce dernier texte, généralement appelé VIM, est disponible en ligne gratuitement sur le site du Bureau International des poids et mesures (BIPM) : <http://www.bipm.org/>. Il est en français et anglais, ce qui est très pratique.

1.1 Unités

1.1.1 Le système international

L'harmonisation de la mesure au niveau international a débuté avec la normalisation des unités à la fin du dix-neuvième siècle. Je ne vais pas entrer dans le détail car les unités du système international (SI) sont en général bien connues des étudiants.

Il est bien entendu impératif de n'utiliser que les unités du système international ou celles dont l'usage est accepté avec le SI, comme par exemple, la minute, l'heure, le jour pour le temps, le litre pour les volumes ou la tonne pour la masse. On peut aussi citer l'électron-volt ou l'unité astronomique en physique.

Pour en savoir plus sur le sujet, je vous recommande la lecture de la brochure *Le système international d'unités* de 2006 du Bureau International des poids et mesures qui peut être téléchargée sur son site : <http://www.bipm.org/>

1.1.2 Symboles

Les seuls symboles autorisés sont ceux du système international. Toutes les notations exotiques que l'on voit parfois, comme *ml* pour *mètre linéaire* sont à proscrire.

Une petite règle grammaticale mérite d'être signalée : il est interdit d'utiliser des symboles après des nombres écrits en toutes lettres. Exemple : *cinq m* est interdit, on écrira *cinq mètres* ou *5 m*. Avec les bars, cela donne, *5 bar* ou *cinq bars*.

1.2 Ecriture des nombres

Cette écriture doit être sans ambiguïté.

1.2.1 Nombres

Pour faciliter la lecture, les nombres sont écrits par tranche de trois chiffres séparés par un espace, mais jamais par une virgule ou un point. Les années, millésimes... font exception à cette règle. En revanche, en français, la **virgule** est obligatoire pour séparer la partie entière de la partie décimale. En anglais, c'est le point.

Exemples

- 3 856 256,025 3
- 8 décembre 2005

Pour les grands nombres, on évitera la règle française $10^{6N} = N$ illion à partir de 10^{12} (10^{12} =billion, 10^{18} =trillion, ...) car il y a risque de confusion avec la règle en usage aux Etats-Unis : $10^{3N} = (N - 1)$ illion (10^9 =billion, 10^{12} = trillion, ...). On préférera les puissances de 10 ou les préfixes méga-, giga-, ... Voir le tableau 1.1. Les préfixes zetta, zepto, yotta et yocto ont été inclus au système international en 1991.

Pour les puissances de 10, l'écriture normale est $1,5 \times 10^8$, mais $1,5.10^8$ est toléré.

Facteur	Nom	Symbole	Facteur	Nom	Symbole
10^1	déca	da	10^{-1}	déci	d
10^2	hecto	h	10^{-2}	centi	c
10^3	kilo	k	10^{-3}	milli	c
10^6	méga	M	10^{-6}	micro	μ
10^9	giga	G	10^{-9}	nano	n
10^{12}	téra	T	10^{-12}	pico	p
10^{15}	péta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y

TAB. 1.1 – Préfixes du système international.

1.2.2 Règles d'écriture et d'arrondi

Les valeurs numériques ne doivent pas être données avec un nombre excessif de chiffres. Pour l'incertitude type ou élargie, deux chiffres significatifs suffisent. Pour la valeur numérique du résultat, *le dernier chiffre à retenir est celui qui a la même position que le deuxième chiffre significatif dans l'expression de l'incertitude.*

L'arrondissement consiste à remplacer un nombre par un autre nombre appelé **arrondi** choisi comme étant le multiple entier le plus proche du nombre donné.

Exemple 1 Avec un intervalle d'arrondissement de 0,1, 12,235 devient 12,2, et 12,251 devient 12,3.

Ce pose alors le problème du cas où le nombre donné finit par 50. La norme NF X 02-001 propose de choisir le multiple entier pair comme arrondi de façon à minimiser l'erreur d'arrondissement quand on traite de longues séries. Des variantes sont possibles.

Exemple 2 Avec un intervalle d'arrondissement de 0,1, 12,25 devient 12,2, et 12,35 devient 12,4. Avec un intervalle d'arrondissement de 1, 109,50 devient 110 et 108,50 devient 108.

Il est toujours recommandé d'arrondir en une seule fois, l'arrondissement en plusieurs étapes pouvant conduire à des erreurs.

Si après un arrondissement le dernier chiffre est un 0, il faut le garder : 60 et 6×10^1 ne sont pas équivalents. 60 a deux chiffres significatifs, alors que 6×10^1 n'en a qu'un. 60 est équivalent à $6,0 \times 10^1$.

1.3 Vocabulaire

Le vocabulaire est aussi codifié, avec sa traduction en anglais. Je ne vais pas reprendre tous les mots, mais retenir ceux qui vont nous obliger à fixer les idées. Je vous renvoie au VIM [2] pour une liste plus complète.

Le premier de la liste est **grandeur (mesurable)** [(measurable) quantity] dont la définition officielle est *attribut d'un phénomène, d'un corps ou d'une substance, qui est susceptible d'être distingué qualitativement et déterminé quantitativement*. Avec des exemples, c'est plus clair : longueur, temps, température. . . d'une manière générale, ou plus particulièrement, longueur d'une tige donnée, . . .

La **dimension d'une grandeur** [dimension of a quantity] est à différencier de l'**unité (de mesure)** [unit (of measurement)].

La **valeur (d'une grandeur)** [value (of a quantity)] est bien entendu *l'expression quantitative d'une grandeur particulière*. La **valeur vraie** [true value] est la valeur que l'on obtiendrait par un mesurage parfait. Elle est par nature indéterminée.

Le **mesurage** [measurement] est *l'ensemble d'opérations ayant pour but de déterminer une valeur d'une grandeur*. C'est une notion concrète qui se distingue de la mesure qui reste abstraite. On parlera ainsi de **principe** [principle] ou **méthode de mesure** [method of measurement] mais du **résultat d'un mesurage** [result of a measurement]. On distingue aussi un **résultat brut** [uncorrected result] d'un **résultat corrigé** [corrected result].

Le **mesurande** [mesurand] est une *grandeur particulière soumise à mesurage*. Par exemple, la célérité du son dans l'air sec de composition (fraction molaire) $N_2 = 0,780\ 8$, $O_2 = 0,209\ 5$, $Ar = 0,009\ 35$ et $CO_2 = 0,000\ 35$ à la température $T = 273,15\ K$ et à la pression $P = 101\ 325\ Pa$. C'est beaucoup plus complet que la grandeur. Ainsi, avant d'effectuer un mesurage, la première étape consiste à spécifier le mesurande. En principe, le mesurande ne pourrait être complètement décrit qu'avec une quantité infinie d'information. Les lacunes en information peuvent parfois introduire une composante dans l'incertitude d'un mesurage.

L'**incertitude de mesure** [uncertainty of measurement] est différente de l'**erreur de mesure** [error of measurement]. Cette dernière, qui est le *résultat d'un mesurage moins une valeur vraie du mesurande*, est, par définition, inaccessible. L'incertitude de mesure *caractérise la dispersion des valeurs qui pourraient raisonnablement être attribuées au mesurande*, et découle, quant à elle, d'un point de vue opérationnel.

Il existe aussi tout un vocabulaire codifié pour les instruments de mesure et leurs caractéristiques, ainsi que leurs étalons. Je n'en parlerai pas.

1.4 Le concept d'incertitude

Dans le passé le concept d'incertitude se confondait avec celui de précision. Maintenant ce dernier terme est à proscrire. Dans la pratique, on effectuait des *calculs d'erreur* maximales qui se combinaient linéairement dans une logique de prudence afin d'encadrer la valeur vraie. De nos jours, on tente plutôt de caractériser la dispersion des valeurs avec une méthode universelle reconnue au niveau international. Cette information est essentielle aux utilisateurs du résultat pour estimer sa fiabilité et donc caractériser la qualité d'un mesurage, voire d'un laboratoire. . .

1.4.1 Historique

La nécessité d'harmoniser les pratiques d'estimation de l'incertitude date de la fin des années 1970. Sous l'égide du Bureau International des Poids et des Mesures (BIPM), une première recommandation (Recommandation INC-1 de 1980) allait être à l'origine de la nouvelle démarche. La publication en 1993 du *Guide to the expression of uncertainty in measurement* communément appelé GUM par le BIPM, la Commission Electrotechnique Internationale (CEI), la Fédération Internationale de Chimie Clinique (FICC), l'Organisation Internationale de Normalisation (ISO), l'Union Internationale de Chimie Pure et Appliquée (UICPA), l'Union Internationale de Physique Pure et Appliquée (UIPPA) et l'Organisation Internationale de Métrologie Légale (OIML), est un pas important. Ce texte est l'un des textes de référence utilisé par quasiment tous les pays du monde. Sa traduction française date de 1995. Il sera repris sous forme de norme européenne et donc française sous la référence NF ENV 13005 en 1999.

C'est cette norme qui va servir de base au chapitre suivant. Un historique plus détaillé est en annexe A.

En 2008, un premier supplément au GUM a été publié sur la *propagation de distributions par une méthode de Monte Carlo* [4]. Ce sera l'objet du chapitre 3.

1.4.2 Les processus de mesure et les processus d'essai

La notion de mesure a été définie précédemment. La métrologie s'y rapporte. Le monde des essais est complémentaire dans le sens où il utilise de nombreuses mesures, dont les grandeurs sont liées pour quantifier une caractéristique donnée. Ce peut être le cas par exemple de la corrosion, de la déformation, de l'étanchéité, de la rigidité. . .

Un **essai** [test] est, selon la norme NF EN ISO CEI 17025, défini comme *l'opération technique qui consiste à déterminer une ou plusieurs caractéristique ou la performance d'un produit, matériau, équipement, organisme, phénomène physique, processus ou service donné, selon un mode opératoire spécifié.*

Il est apparu que la démarche du GUM était parfois difficile voire impossible à mettre en oeuvre dans le cadre des essais. La quantification de la qualité d'une méthode d'essai est décrite dans la norme NF ISO 5725 complétée du fascicule de documentation FD X 07-021 *Aide à la démarche pour l'estimation et l'utilisation de l'incertitude des mesures et des résultats d'essais.* Ces textes sont complémentaires du GUM, mais dépassent le cadre de ce cours qui se limitera aux mesures.

1.4.3 Facteurs d'influence

Tout mesurage est entaché d'erreurs qui ont pour origine, le système de mesure et le mesurande. En effet, un système de mesure n'est jamais parfait. Il peut être sensible à l'environnement et dépend des étalons qui ont servi à son étalonnage. Quand on répète les observations, on observe souvent une dispersion des valeurs qui montre que le système n'est pas fidèle. Le mesurande peut être mal défini, ce qui est source d'incertitude.

Analyser correctement le processus de mesure est certainement l'étape la plus difficile de l'estimation de l'incertitude et nécessite de bien maîtriser la technique de mesure. Pour recenser tous les **facteurs d'influence** [influence factor] pouvant jouer a priori sur l'incertitude, il existe une technique dite des 5M basée sur un diagramme dit d'Ishikawa.

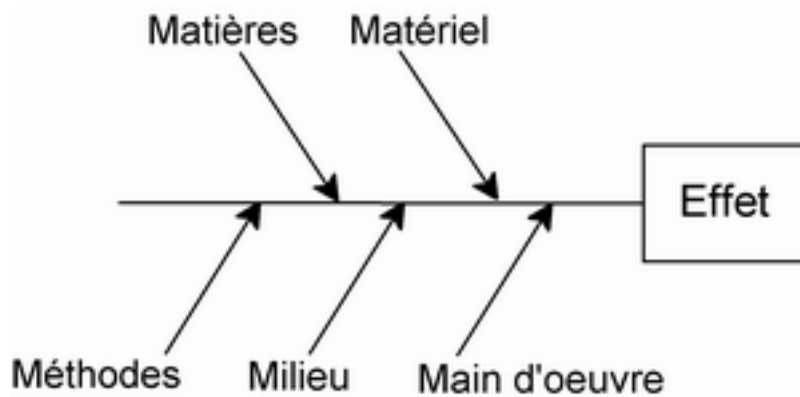


FIG. 1.1 – Diagramme d'Ishikawa pour la méthode des 5M.

La méthode des cinq M consiste à rechercher les facteurs d'influence suivant 5 rubriques et à les classer sur un diagramme d'Ishikawa¹ selon cinq rubriques explicitées sur la figure 1.1. Dans chaque rubrique, des flèches supplémentaires font apparaître les causes et éventuellement des sous-causes. Ce travail est aussi utile pour corriger le processus de mesure et réduire l'incertitude.

Voici quelques exemples pour chaque rubrique :

Main d'oeuvre Prélèvements

Méthode Mode opératoire de l'analyse, nombre de répétitions

¹Le professeur Kaoru Ishikawa (1915-1989) est à la base de plusieurs méthodologies de la démarche qualité.

Milieu Concerne les facteurs environnementaux lors de prélèvements ou lors du mesurage, comme la température.

Matériel Le matériel et les instruments sont source d'incertitude

Matière Concerne la qualité de l'échantillon à analyser. Est-il bien représentatif?

Une fois ce travail effectué, il faut pouvoir estimer les incertitudes liées à chaque facteur d'influence (quand elles ne sont pas données) et apprendre à les combiner pour finalement estimer l'incertitude sur le résultat du mesurage. C'est l'objet du chapitre suivant.

Chapitre 2

Estimation de l'incertitude suivant la méthode du GUM

Ce chapitre est basé sur le GUM, qui a le statut de prénorme européenne et de norme française NF ENV 13005 : **Guide pour l'expression de l'incertitude de mesure**, août 1999. Des versions française et anglaise sont disponibles gratuitement sur le site du BIPM : <http://www.bipm.org/>. Pour la chimie, il existe un site en anglais qui publie plusieurs guides sur le calcul d'incertitude : <http://www.measurementuncertainty.org/>

Pour ce chapitre, je me suis servi des ouvrages suivants :

- *Erreurs de mesure*, par Michèle Neuilly, techniques de l'ingénieur, traité Mesures et Contrôles.
- *Modélisation et estimation des erreurs de mesure*, par Michèle Neuilly, Tec&Doc 1998.
- *Estimer l'incertitude. Mesures - Essais*, Christophe Perruchet et Marc Priel, AFNOR 2000.
- *An introduction to error analysis, the study of uncertainties in physical environments* John R. Taylor, University Science Books, (1982)

Il en existe de nombreux autres.

Nous allons d'abord traduire en langage mathématique les définitions vues dans le chapitre précédent, puis faire quelques rappels de statistique avant d'aborder les calculs d'incertitude normalisés et l'expression des résultats.

2.1 Quand le résultat de mesure est une variable aléatoire

Quand on cherche à estimer une grandeur X , les résultats de mesurages successifs sont dits *aléatoires* quand ils varient de façon imprévisible. Un premier mesurage peut donner, par exemple, $x_1 = 724,173$ g, puis après avoir recommencé toute la procédure à partir du tarage on peut obtenir $x_2 = 724,190$ g... Comment estimer la valeur vraie ?

On peut faire un histogramme des effectifs pour présenter les résultats, voir figure 2.1. En ordonnée, on met n_{ab} le nombre de valeurs comprises dans l'intervalle $[a, b]$. n_{ab}/n , où n est le nombre total de valeurs, est appelé *fréquence*. Cette série de valeurs peut être caractérisée statistiquement par sa moyenne arithmétique,

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad (2.1)$$

et un écart quadratique moyen,

$$d = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}}. \quad (2.2)$$

Quand le nombre d'observations augmente infiniment et que la sensibilité de l'instrument de mesure peut aussi être augmentée à l'infini, l'histogramme normalisé au nombre total n tend vers une courbe continue appelée **courbe de distribution**. Evidemment, cette courbe est inaccessible, mais l'objet de la statistique est d'estimer ses caractéristiques (forme et position) à partir d'une série limitée de mesurages.

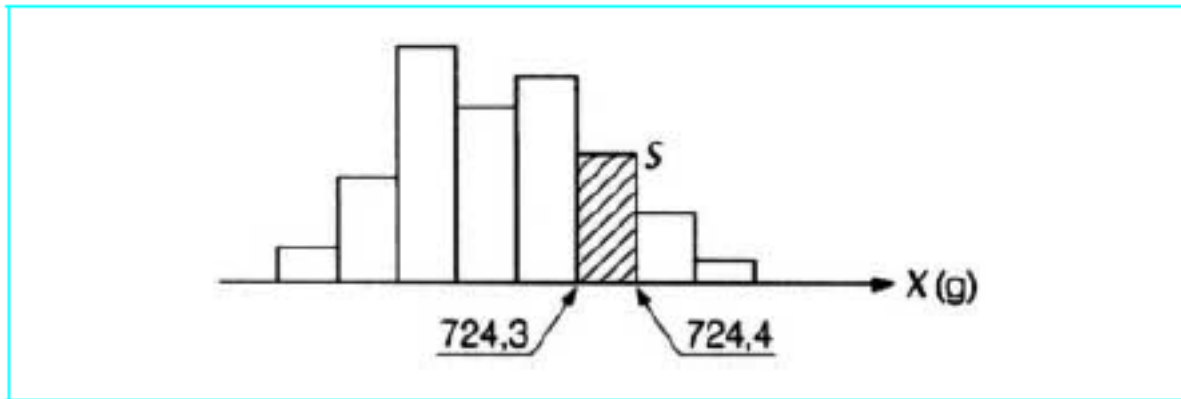


FIG. 2.1 – Histogramme des effectifs des valeurs issues d’une série de mesurages (figure tirée de M. Neuilly)

Quand $n \rightarrow \infty$, on admet que \bar{x} tend vers une valeur limite, μ , qui est appelée **moyenne de la population** [sample mean] ou **valeur probable de x** ou **espérance mathématique** [expectation] $E(X)$,

$$\mu = E(X) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{x_i}{n}. \quad (2.3)$$

Si la courbe de distribution est symétrique, il s’agit du maximum. La dispersion des résultats est caractérisée par la **variance** [variance] σ^2 ,

$$\text{Var}(X) = \sigma^2 = E(X - \mu)^2 = \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n}. \quad (2.4)$$

La racine carrée de la variance, σ est appelée **écart-type** [standard deviation]. On peut montrer que

$$\text{Var}(X) = \sigma^2 = E(x^2) - E^2(x) = E(x^2) - \mu^2 = \lim_{n \rightarrow \infty} \left[\sum_{i=1}^n \frac{x_i^2}{n} - \bar{x}^2 \right]. \quad (2.5)$$

Les lois de probabilité sont généralement inconnues, mais il existe des modèles de distributions qui peuvent correspondre à nombre de situations expérimentales. La loi la plus utilisée est la **loi normale** ou **loi de Gauss**, ou **Laplace-Gauss**,

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp - \left[\frac{(x - \mu)^2}{2\sigma^2} \right], \quad (2.6)$$

où μ et σ sont des constantes caractéristiques. L’intérêt pour cette loi réside dans le fait qu’un résultat de mesure suit une loi normale si les conditions suivantes sont réalisées simultanément :

- les causes d’erreur sont nombreuses ;
- les erreurs sont du même ordre de grandeur ;
- les fluctuations liées aux différentes causes d’erreur sont indépendantes et additives.

C’est une conséquence du théorème de la limite centrale.

Si la courbe de distribution suit une loi normale, il est aisé de connaître la probabilité d’obtenir une valeur donnée en fonction de σ , voir figure 2.2. Dans le cas a, l’intervalle $[\mu - \sigma, \mu + \sigma]$ inclut 68% des résultats. Les valeurs avec 2σ et 3σ sont données figure 2.2. 99% correspond à $\pm 2,58\sigma$. La largeur à mi-hauteur [full width at half maximum, FWHM] est parfois utilisée et vaut $FWHM = 2,36\sigma$. L’écart-type σ permet de caractériser la dispersion des résultats et sert donc à estimer les incertitudes.

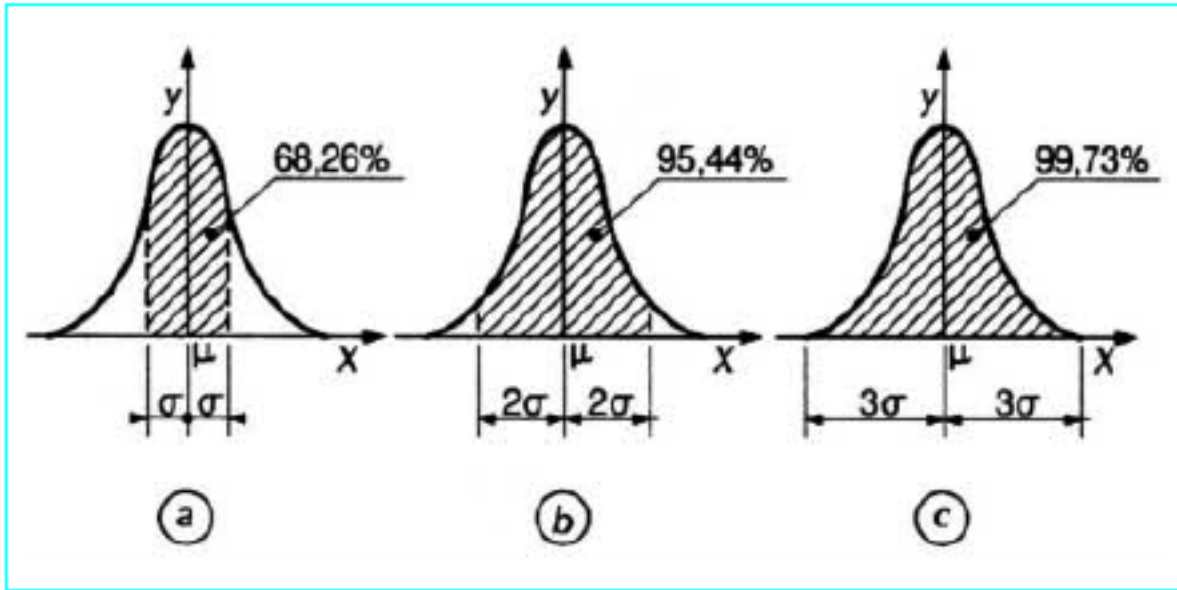


FIG. 2.2 – Distribution normale et largeurs (figure tirée de M. Neuilly)

2.2 Détermination expérimentale de la variance : évaluation de type A de l'incertitude

Le GUM définit l'évaluation de type A de l'incertitude [type A evaluation of measurement uncertainty], comme étant "l'évaluation d'une composante de l'incertitude de mesure par une analyse statistique des valeurs mesurées obtenues dans des conditions définies de mesurage".

2.2.1 Estimateur de la variance

Définition

Pour accéder à la variance, il faut faire une infinité d'observations indépendantes infiniment précises, ce qui est bien entendu inaccessible. En revanche, on peut estimer cette variance en répétant les observations dans les mêmes conditions bien définies.

Soit $\{x_1, \dots, x_n\}$ les résultats obtenus et

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} \quad (2.7)$$

leur moyenne arithmétique. On peut montrer que la variance peut être estimée par

$$s^2(x_i) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.8)$$

On prend $n - 1$ degrés de liberté car il y a n valeurs moins une corrélation (moyenne). On peut aussi utiliser

$$s^2(x_i) = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right] = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - n(\bar{x})^2 \right]. \quad (2.9)$$

Pour obtenir cette forme, il suffit de développer le carré dans l'équation (2.8).

Variance de la moyenne

Si X est la somme de deux grandeurs indépendantes A et B : $X = A + B$, alors $Var(X) = Var(A) + Var(B)$. En effet,

$$Var(X) = \lim_{N \rightarrow \infty} \sum_{k=1}^N [a_k + b_k - (\mu_a + \mu_b)]^2 = \lim_{N \rightarrow \infty} \sum_{k=1}^N [(a_k - \mu_a)^2 + (b_k - \mu_b)^2 + 2(a_k - \mu_a)(b_k - \mu_b)]. \quad (2.10)$$

Pour $N \rightarrow \infty$, les termes croisés s'annulent en moyenne si les a_k et b_k sont indépendants et l'on a finalement cqfd. C'est la formule de Bienaymé.

On en déduit que

$$Var(\bar{x}) = Var\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \sum_{i=1}^n Var\left(\frac{x_i}{n}\right) = \frac{1}{n^2} \sum_{i=1}^n Var(x_i) = \frac{1}{n} Var(x_i). \quad (2.11)$$

On distinguera donc bien la variance associée à une valeur de mesure $Var(x_i)$ de la variance associée à la valeur moyenne de ces valeurs, $Var(\bar{x}) = \frac{1}{n} Var(x_i)$. Plus le nombre de mesurages est grand plus la variance sera petite pour la valeur moyenne.

Quand on a une série de valeurs $\{x_1, \dots, x_n\}$, on n'a, a priori, pas de préférence pour une valeur plutôt qu'une autre. Si on en prend une au hasard comme estimation de la valeur du mesurande, l'incertitude associée sera $s(x_i)$. En revanche, si l'on choisit la moyenne pour estimer la valeur du mesurande, l'incertitude associée, $s(\bar{x})$ est plus petite. C'est pourquoi on utilise la moyenne.

Finalement, la variance de la moyenne des valeurs des mesurages peut être estimée par

$$s^2(\bar{x}) = \frac{1}{n} s^2(x_i) = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.12)$$

Estimateur non biaisé [unbiased estimator]

Pourquoi pour estimer la variance, on divise $n - 1$ et non n comme on aurait pu le penser naïvement ? Supposons que l'espérance et la variance $Var(X) = \sigma^2$ soient estimées pour une première série de n mesurages par

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{et} \quad s^2(x_i) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (2.13)$$

Si l'on fait une autre série de n mesurages, on obtiendrait une autre valeur pour \bar{x} et $s^2(x_i)$.

La moyenne sur un nombre infini de séries des moyennes estimées \bar{x} donne bien μ . Le calcul est immédiat. Cet estimateur n'est pas biaisé.

Le calcul de la moyenne sur un nombre infini de séries des variances estimées est un peu plus long :

$$\begin{aligned} E(s^2(x_i)) &= E\left(\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}\right) = \frac{1}{n-1} E\left(\sum_{i=1}^n (x_i - \mu + \mu - \bar{x})^2\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^n (x_i - \mu)^2 - n(\bar{x} - \mu)^2\right) = \frac{1}{n-1} \left(\sum_{i=1}^n E((x_i - \mu)^2) - nE((\bar{x} - \mu)^2)\right) \end{aligned} \quad (2.14)$$

Le premier terme $E((x_i - \mu)^2) = Var(X) = \sigma^2$ et le deuxième $E((\bar{x} - \mu)^2) = Var(\bar{x}) = \sigma^2/n$. Ainsi,

$$E(s^2(x_i)) = \frac{1}{n-1} \left[\sum_{i=1}^n \sigma^2 - n(\sigma^2/n)\right] = \frac{n-1}{n-1} \sigma^2. \quad (2.15)$$

Avec n au dénominateur, la moyenne des variances estimées ne donnerait pas σ^2 mais $\frac{n-1}{n} \sigma^2$. L'estimateur serait dit **biaisé** [biased estimator].

Cette correction est due à Friedrich Wilhelm Bessel et est appelée correction de Bessel. Voir par exemple *Bessel's correction* sur wikipedia en anglais.

En plus d'estimer la moyenne des $s^2(x_i)$, on peut aussi calculer leur variance. On obtient que l'incertitude relative sur $s^2(x_i)$ est égale à $1/\sqrt{2(n-1)}$. Le calcul est fait dans le livre de Taylor [6] par exemple. L'annexe E du GUM propose une formule plus complexe basée sur la distribution du χ^2 . Je ne vais pas entrer dans le détail, mais juste reprendre quelques valeurs numériques dans le tableau 2.1.

nombre d'observations	incertitude relative sur $s(\bar{x})$
2	76%
3	52%
4	42%
5	36%
10	24%
20	16%
30	13%
50	10%

TAB. 2.1 – Incertitude relative sur $s(\bar{x})$ en fonction du nombre d'observations.

2.2.2 Mise en œuvre pratique

- Lors de la saisie, il faut veiller à ce que les différences $x_i - \bar{x}$ aient au moins deux chiffres significatifs.
- Il est possible de retrancher une valeur fixe de façon à alléger les calculs. Supposons, par exemple, que $x_1 = 1,799\ 50, x_2 = 1,799\ 42, x_3 = 1,799\ 63, \dots$. Il est plus pratique d'utiliser $y_i = x_i - 1,799\ 00$. On a alors aisément que

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2. \quad (2.16)$$

- Si on utilise deux méthodes de mesure indépendantes donnant chacune une série de valeurs caractérisée par \bar{x}_1 et s_1^2 , et \bar{x}_2 et s_2^2 , il est possible de combiner ces deux séries

$$\bar{x} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2} \quad (2.17)$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}, \quad (2.18)$$

où n_i est le nombre de valeurs de chaque série. On peut généraliser à un plus grand nombre de séries.

Combien faut-il de valeurs pour avoir une estimation raisonnable de la variance ? Intuitivement, pour couvrir toute la gamme de mesure qui a une largeur 6σ , on se dit qu'il faut au moins 6 mesures. Le tableau 2.1 donne une estimation de l'incertitude relative de la variance estimée par rapport à la variance vraie en fonction du nombre d'observations.

Au début, l'incertitude relative varie vite quand on augmente le nombre d'observations. Puis, à partir de 10 environ, il en faut beaucoup plus pour faire baisser de manière significative cette incertitude. Une dizaine de valeurs constitue alors un bon compromis.

Nous avons fait implicitement l'hypothèse ici que les résultats $\{x_1, \dots, x_n\}$ étaient connus avec une précision infinie, sans incertitude de lecture par exemple, ce qui n'est pas réaliste. L'évaluation de l'incertitude en répétant les mesures devra être combinée avec les autres sources d'incertitudes. Il faut donc bien définir la frontière entre ce qui est pris en compte dans un mesurage et ce qui doit être ajouté par ailleurs. Ainsi, la lecture successive d'un baromètre ne va pas changer si l'on ne le déséquilibre pas à chaque mesure. Plus généralement, si la mise à zéro fait partie du mode opératoire, il faut l'effectuer à chaque fois. La question à se poser est : "Dans quelle mesure les observations répétées sont-elles bien indépendantes entre elles ?". Il faut ensuite bien préciser le mesurande lorsque l'on affiche les résultats. Si l'on analyse toujours le même échantillon, on aura des informations sur cet échantillon, et non sur le matériau analysé. Les autres contributions à l'incertitude devront être incorporées à l'aide d'un calcul d'incertitude.

2.3 Evaluation de type B de l'incertitude

L'évaluation de type B de l'incertitude [type B evaluation of measurement uncertainty] est définie très simplement : "évaluation d'une composante de l'incertitude de mesure par d'autres moyens qu'une évaluation de type A de l'incertitude".

L'écart-type issu d'une évaluation de type A est généralement noté s et celui d'une évaluation de type B, u .

Les sources d'information pour estimer les incertitudes-types de type B sont multiples :

- résultats de mesures antérieures,
- spécifications du fabricant,
- valeurs de référence provenant de tables ou manuels...

Ces données sont parfois incomplètes et il faut faire appel à sa perspicacité. Le résultat peut être aussi fiable que pour les incertitudes de type A, surtout quand ces dernières sont obtenues à partir d'un petit nombre de mesures. Nous allons traiter un petit nombre de cas typiques.

- Le fabricant, au fait des normes, donne directement une incertitude type. Il n'y a donc aucun problème.
- Le fabricant donne une incertitude avec un niveau de confiance (90, 95 ou 99%). Sauf indication contraire, on peut penser qu'il a utilisé une loi normale et l'on peut retrouver facilement l'incertitude type en divisant respectivement par le facteur correspondant.
- La loi de probabilité est parfois prise rectangulaire faute de mieux. On a alors une probabilité constante d'être dans un intervalle $[a_-, a_+]$ et 0 en dehors. C'est le cas en particulier quand on lit une valeur sur une échelle graduée, où l'intervalle correspond à deux graduations successives. Dans ces conditions, l'espérance est prise au milieu de l'intervalle, $\mu = \frac{a_+ + a_-}{2}$ et

$$u^2 = \frac{(a_+ - a_-)^2}{12} = \frac{a^2}{3}, \quad (2.19)$$

si l'on note $2a$ la longueur de l'intervalle. Si la moyenne n'est pas au centre de l'intervalle, ce qui suppose que la loi de probabilité n'est pas symétrique, alors on prend encore ce type de règle, faute de mieux.

- Pour une loi trapézoïdale ayant une base de longueur $2a$ et un sommet de largeur $2\beta a$,

$$u^2 = \frac{a^2(1 + \beta^2)}{6}. \quad (2.20)$$

Si $\beta = 1$ on retrouve le cas de la loi rectangulaire. Si $\beta = 0$ on obtient celle du triangle $u^2 = a^2/6$.

Le GUM liste de nombreux autres cas. D'une manière générale, l'espérance et la variance sont évaluées à partir des lois de distribution choisies par

$$\mu = \frac{\int_{-\infty}^{\infty} xP(x) dx}{\int_{-\infty}^{\infty} P(x) dx} \quad \text{et} \quad \sigma^2 = \frac{\int_{-\infty}^{\infty} (x - \mu)^2 P(x) dx}{\int_{-\infty}^{\infty} P(x) dx}. \quad (2.21)$$

2.4 Propagation de l'incertitude

On mesure rarement directement la grandeur qui nous intéresse et l'on doit passer par des opérations mathématiques qui incluent souvent d'autres grandeurs affectées d'une incertitude. Par exemple, une résistance électrique peut être évaluée en mesurant U et I , puis en calculant U/I . La méthode préconisée par le GUM pour évaluer la propagation de l'incertitude diffère de celle qui est généralement enseignée à l'université et nécessite donc des explications.

2.4.1 Méthode

Supposons que la **grandeur de sortie dans un modèle de mesure** [output quantity in measurement model] Z dépende de N **grandeurs d'entrée** [input quantities] w_1, \dots, w_N décrite par une loi de probabilité : $Z = f(w_1, \dots, w_N)$. Si l'on note $\mu_z = f(\bar{w}_1, \dots, \bar{w}_N)$ l'espérance, l'écart au premier ordre est

$$z - \mu_z = \sum_{i=1}^N \frac{\partial f}{\partial w_i} (w_i - \bar{w}_i). \quad (2.22)$$

En élevant au carré, on obtient,

$$(z - \mu_z)^2 = \sum_{i=1}^N \left(\frac{\partial f}{\partial w_i} \right)^2 (w_i - \bar{w}_i)^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial w_i} \frac{\partial f}{\partial w_j} (w_i - \bar{w}_i)(w_j - \bar{w}_j). \quad (2.23)$$

En prenant la valeur moyenne, on obtient finalement,

$$\sigma_z^2 = \sum_{i=1}^N \left(\frac{\partial f}{\partial w_i} \right)^2 \sigma^2(w_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial f}{\partial w_i} \frac{\partial f}{\partial w_j} \sigma(w_i, w_j). \quad (2.24)$$

Le deuxième terme prend en compte les corrélations éventuelles entre les w_i via la **covariance** [covariance] $\sigma(w_i, w_j)$. Nous allons étudier ces corrélations dans la suite. Nous nous intéressons d'abord au cas de grandeurs w_i indépendantes pour lesquelles seul le premier terme est à prendre en considération.

Cette méthode consiste à évaluer l'**incertitude type** [standard uncertainty] qui est l'incertitude du résultat d'un mesurage exprimée sous la forme d'un écart-type.

2.4.2 Exemples avec des grandeurs indépendantes

Exemple 1

Comparons cette méthode à la méthode traditionnelle sur l'exemple simple d'une résistance mesurée à partir de l'intensité I et de la différence de potentiel U . L'ancienne méthode utilise la différentielle logarithmique,

$$\frac{dR}{R} = \frac{dU}{U} - \frac{dI}{I}, \quad (2.25)$$

pour encadrer l'incertitude en considérant le cas le plus pessimiste,

$$\frac{\Delta R}{R} = \frac{\Delta U}{U} + \frac{\Delta I}{I}. \quad (2.26)$$

La méthode préconisée donne

$$\sigma^2(R) = \frac{\sigma^2(U)}{I^2} + \frac{U^2}{I^4} \sigma^2(I), \quad (2.27)$$

ou en relatif,

$$\frac{\sigma(R)}{R} = \sqrt{\frac{\sigma^2(U)}{U^2} + \frac{\sigma^2(I)}{I^2}}. \quad (2.28)$$

Une simple application numérique supposant que U et I sont connues à 10% près, donne

$$\frac{\Delta R}{R} = 20\% \quad \text{et} \quad \frac{\sigma(R)}{R} = 14\%. \quad (2.29)$$

Cette différence de résultats s'explique par la différence de philosophie de chacune des approches. Dans le premier cas, il s'agit d'encadrer une valeur en étant pessimiste, alors que dans le deuxième, il s'agit plutôt d'estimer une dispersion des valeurs qui pourraient être raisonnablement attribuées au mesurande pour ensuite attribuer un degré de confiance au résultat. La démarche suivie se veut opérationnelle.

Exemple 2

Considérons le cas d'un calibrage avec une simple loi affine entre deux grandeurs, $Z = \alpha + \beta Q$, où α et β sont aussi supposées être caractérisées par des lois de probabilité données par le constructeur de l'appareil. Ainsi, pour chaque mesurage, on a

$$z_k = \alpha + \beta q_k. \quad (2.30)$$

On s'intéresse à la valeur moyenne de Z ,

$$\bar{z} = \frac{1}{n} \sum_{k=1}^n z_k = \frac{1}{n} \sum_{k=1}^n (\alpha + \beta q_k) = \alpha + \beta \bar{q}, \quad (2.31)$$

que l'on considère être la meilleure estimation de z . Pour estimer l'incertitude-type, on part de $\bar{z} = f(\alpha, \beta, q_k)$ pour obtenir

$$\sigma^2(\bar{z}) = u^2(\alpha) + \bar{q}^2 u^2(\beta) + \sum_{k=1}^n \left(\frac{\beta}{n}\right)^2 s^2(q_k) \quad (2.32)$$

$$= u^2(\alpha) + \bar{q}^2 u^2(\beta) + \beta^2 \frac{s^2(q_k)}{n}. \quad (2.33)$$

Dans cette expression, nous avons distingué la variance venant directement d'une série de mesures, notée $s^2(q_k)$, des variances venant d'autres sources, notées u^2 . Ici $u^2(\alpha)$ et $u^2(\beta)$ sont données par le fabricant. Dans la terminologie traditionnelle, le troisième terme est appelé **contribution aléatoire** et décroît lorsque n augmente. Les deux premiers termes sont appelés **contributions systématiques** car ils ne dépendent pas de n . Par ce traitement, on ne fait plus de distinction entre les effets systématiques et les effets aléatoires.

Dans cet exemple, on peut considérer le cas particulier où $\alpha = 0$ et $\beta = 1$ qui conduit à

$$s^2(\bar{q}) = \frac{1}{n} s^2(q_k). \quad (2.34)$$

On retrouve que la variance sur une moyenne diminue donc avec le nombre de valeurs. D'où l'intérêt d'augmenter le nombre de mesurages. Evidemment, cette relation n'est valable que si les répétitions sont bien indépendantes entre elles.

Exercices :

- En supposant que u et v sont deux grandeurs indépendantes, trouver la relation entre l'incertitude sur x , σ_x en fonction des incertitudes σ_u et σ_v pour les fonctions suivantes :
 - $x = \frac{1}{2}(u + v)$
 - $x = \frac{1}{2}(u - v)$
 - $x = u^2 + v^2$
 - $x = 1/u^2$
 - $x = uv^2$
 - $x = u^n v^m$.

Pour les trois derniers exemples, calculer aussi l'incertitude relative $\sigma_x/|x|$.

- Si le diamètre d'une table ronde est connu à 1% près, quelle est l'incertitude sur sa surface ? Est-il préférable de connaître son rayon à 1% près ?
- L'activité initiale N_0 et la demi-vie τ d'une source radioactive sont connues avec une incertitude relative de 1% chacune. Estimer l'incertitude sur l'activité à la date t , $N(t) = N_0 e^{-t/\tau}$. Pour quelle valeur de t/τ les incertitudes sur N_0 et τ contribuent-elles autant à l'incertitude sur $N(t)$?

2.4.3 Prise en compte des corrélations éventuelles

Il peut y avoir une corrélation significative entre deux grandeurs d'entrée si l'on utilise pour leur détermination,

- le même instrument de mesure,
- le même étalon physique, ou
- la même donnée de référence ayant une incertitude-type significative.

Cette corrélation peut être mesurée à l'aide de la covariance, $\sigma(w_i, w_j)$, ou plus généralement, à l'aide du **coefficient de corrélation** [correlation coefficient],

$$r(w_i, w_j) = \frac{\sigma(w_i, w_j)}{\sigma(w_i)\sigma(w_j)}, \quad (2.35)$$

qui est toujours compris entre -1 et 1. Evidemment, si w_i et w_j sont indépendants, $r(w_i, w_j) = 0$, alors que si $r = \pm 1$, il existe une relation linéaire entre w_i et w_j de la forme $w_j = Aw_i + B$.

A noter que la covariance a la dimension d'une variance (σ^2) et est donc notée $\sigma^2(w_i, w_j)$ dans certains livres. Mais cette fonction peut être négative et donc, nous garderons l'usage du GUM.

Ainsi, il y a corrélation quand un doute commun pèse sur deux estimations w_i, w_j . Pour chercher à savoir si deux grandeurs d'entrée sont corrélées ou non, il ne faut pas raisonner uniquement sur les mesurands correspondants, mais analyser comment le processus de mesure a pu introduire, ou n'a pu éliminer un doute commun à ces deux variables aléatoires qui représentent les mesurands.

La covariance ou le coefficient de corrélation sont estimés soit directement à partir des grandeurs mesurées, soit par calcul à partir d'autres grandeurs.

Evaluation de type A

Soient deux moyennes arithmétiques \bar{a} et \bar{b} qui estiment les espérances mathématiques μ_a et μ_b de deux grandeurs A et B aléatoires. Supposons que \bar{a} et \bar{b} soient calculées à partir de n paires indépendantes d'observations simultanées de A et B faites dans les mêmes conditions de mesure. Alors la covariance peut être estimée par

$$s(a_l, b_l) = \frac{1}{n-1} \sum_{k=1}^n (a_k - \bar{a})(b_k - \bar{b}). \quad (2.36)$$

La covariance des valeurs moyennes, \bar{a} et \bar{b} , est obtenue en divisant par n

$$s(\bar{a}, \bar{b}) = \frac{1}{n} s(a_l, b_l) = \frac{1}{n(n-1)} \sum_{k=1}^n (a_k - \bar{a})(b_k - \bar{b}). \quad (2.37)$$

Le coefficient de corrélation peut alors être aisément calculé en divisant par les écarts-types respectifs.

Estimation indirecte

Si les grandeurs A et B ne sont pas mesurées directement, mais dépendent chacune d'un ensemble de N variables **non corrélées** w_i , la covariance vaut,

$$\sigma(a_l, b_l) = \sum_{i=1}^N \frac{\partial A}{\partial w_i} \frac{\partial B}{\partial w_i} \sigma^2(w_i). \quad (2.38)$$

Pour obtenir le coefficient de corrélation, il faut diviser par les écarts-types obtenus à partir des variances rappelées ici

$$\sigma^2(a_l) = \sum_{i=1}^N \left(\frac{\partial A}{\partial w_i} \right)^2 \sigma^2(w_i) \quad (2.39)$$

$$\sigma^2(b_l) = \sum_{i=1}^N \left(\frac{\partial B}{\partial w_i} \right)^2 \sigma^2(w_i). \quad (2.40)$$

Si les grandeurs d'entrées sont elles-mêmes corrélées, l'expression de la covariance se complique

$$\sigma(a_l, b_l) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial A}{\partial w_i} \frac{\partial B}{\partial w_j} \sigma(w_i, w_j). \quad (2.41)$$

On passe d'une somme de N termes à une somme de N^2 termes.

Exemple 1

Cet exemple est tiré de la norme FD X 07-021

Deux masses m_1 et m_2 de 50 g chacune sont utilisées pour réaliser un assemblage m_a de 100 g, sachant qu'elles ont été elles-mêmes comparées au même étalon m_e de 50 g. Quelle est l'incertitude-type composée de cet assemblage ?

Dans cet exemple simple, on peut exprimer m_a directement à partir des grandeurs mesurées,

$$m_a = m_e + x_1 + m_e + x_2, \quad (2.42)$$

où x_1 et x_2 sont les erreurs dues à la pesée des masses m_1 et m_2 respectivement. Si on suppose que x_1 et x_2 sont indépendantes, on a directement,

$$\sigma^2(m_a) = \sigma^2(x_1) + \sigma^2(x_2) + 4\sigma^2(m_e). \quad (2.43)$$

Mais, si on adopte une méthode par étape que l'on pourrait trouver dans un exemple plus complexe, il faut tenir compte des termes de covariance :

$$m_a = m_1 + m_2, \quad \text{et donc} \quad \sigma^2(m_a) = \sigma^2(m_1) + \sigma^2(m_2) + 2\sigma(m_1, m_2). \quad (2.44)$$

Pour déterminer chaque terme, il faut expliciter mathématiquement la méthode de mesure :

$$m_1 = m_e + x_1 \quad \text{et} \quad m_2 = m_e + x_2. \quad (2.45)$$

Et donc, en appliquant les formules du GUM, on a,

$$\sigma^2(m_1) = \sigma^2(m_e) + \sigma^2(x_1), \quad \sigma^2(m_2) = \sigma^2(m_e) + \sigma^2(x_2) \quad \text{et} \quad \sigma(m_1, m_2) = \frac{\partial m_1}{\partial m_e} \frac{\partial m_2}{\partial m_e} \sigma^2(m_e) = \sigma^2(m_e). \quad (2.46)$$

Finalement,

$$\sigma^2(m_a) = \sigma^2(x_1) + \sigma^2(x_2) + 4\sigma^2(m_e). \quad (2.47)$$

On constate que selon la modélisation choisie, il y a un terme de covariance ou il n'y en a pas. Dans le deuxième cas, si on néglige la covariance, on ne retrouve pas le bon résultat.

Exemple 2

Cet exemple est tiré de la norme NF ENV 13005.

Une même résistance étalon R_s est utilisée pour la mesure

– d'une intensité, $I = V_s/R_s$, et

– d'une température à l'aide d'une résistance, $t = aR^2(t) - t_0$. $R(t)$ est déterminée à l'aide d'un pont dans lequel intervient R_s . On mesure alors $\beta = R(t)/R_s$. Ainsi, $t = a\beta^2 R_s^2 - t_0$.

La covariance entre I et t est facilement estimée car la seule grandeur d'entrée commune est R_s et vaut alors,

$$\sigma(I, t) = \frac{\partial I}{\partial R_s} \frac{\partial t}{\partial R_s} \sigma^2(R_s) = -2 \frac{I(t + t_0)}{R_s^2} \sigma^2(R_s). \quad (2.48)$$

Supposons maintenant que I et t servent à déterminer

$$P = \frac{C_0 I^2}{T_0 + t}, \quad (2.49)$$

où C_0 et T_0 sont des constantes dont on négligera l'incertitude. Le calcul de l'incertitude type relative de P doit alors tenir compte de la corrélation entre I et t ,

$$\frac{\sigma^2(P)}{P^2} = 4 \frac{\sigma^2(I)}{I^2} + \frac{\sigma^2(t)}{(T_0 + t)^2} - 4 \frac{\sigma(I, t)}{I(T_0 + t)}, \quad (2.50)$$

avec

$$\frac{\sigma^2(I)}{I^2} = \frac{\sigma^2(V_s)}{V_s^2} + \frac{\sigma^2(R_s)}{R_s^2} \quad (2.51)$$

$$\frac{\sigma^2(t)}{(T_0 + t)^2} = 4 \frac{\sigma^2(\beta)}{\beta^2} + 4 \frac{\sigma^2(R_s)}{R_s^2}. \quad (2.52)$$

Comme chaque incertitude-type relative sur I et t contient une contribution de celle sur R_s , sans le terme de corrélation pour P , la contribution de R_s serait sur-estimée dans $\sigma(P)$.

Exercice

Mesurage simultané d'une résistance et d'une réactance (d'après la NORME AFNOR NF ENV 13005)

On détermine la résistance R et la réactance X d'un élément de circuit par la mesure simultanée de l'amplitude V d'une différence de potentiel sinusoïdale entre ses bornes, de l'intensité I du courant alternatif qui le traverse et du déphasage ϕ entre la différence de potentiel alternative et le courant alternatif. Il en résulte que les trois grandeurs d'entrée sont V, I et ϕ et que les trois grandeurs de sortie - les mesurandes - sont les trois composantes de l'impédance : R, X et Z :

$$R = \frac{V}{I} \cos \phi, \quad X = \frac{V}{I} \sin \phi, \quad Z = \frac{V}{I}. \quad (2.53)$$

Combien de grandeurs de sortie indépendantes existent dans cet exemple ? On veut calculer R, X et Z , ainsi que leurs variances estimées et leurs coefficients de corrélation estimés, par deux méthodes à partir des données du tableau 2.2.

Numéro de l'ensemble	V (V)	I (mA)	ϕ (rad)
1	5,007	19,663	1,0456
2	4,994	19,639	1,0438
3	5,005	19,640	1,0468
4	4,990	19,685	1,0428
5	4,999	19,678	1,0433

TAB. 2.2 – Valeurs des grandeurs d'entrée V, I et ϕ obtenues à partir de cinq ensembles d'observations simultanées

Comparez les deux approches possibles :

1/ En utilisant les données du tableau 2.2, calculer les grandeurs statistiques associées à V, I et ϕ . Vérifier que l'on trouve

$$\bar{V} = 4,999 \text{ V} \quad \bar{I} = 19,661 \text{ mA} \quad \bar{\phi} = 1,04446 \text{ rad} \quad (2.54)$$

$$s(\bar{V}) = 0,0032 \text{ V} \quad s(\bar{I}) = 0,0095 \text{ mA} \quad s(\bar{\phi}) = 0,00075 \text{ rad} \quad (2.55)$$

$$r(\bar{V}, \bar{I}) = -0,36 \quad r(\bar{V}, \bar{\phi}) = 0,86 \quad r(\bar{I}, \bar{\phi}) = -0,65. \quad (2.56)$$

En déduire les R, X et Z , ainsi que leurs variances estimées et leurs coefficients de corrélation estimés à l'aide des formules du GUM. Montrer que l'on trouve,

$$R = (127,732 \pm 0,071) \Omega \quad X = (219,847 \pm 0,295) \Omega \quad Z = 254,260 \pm 0,236 \Omega \quad (2.57)$$

$$r(R, X) = -0,588 \quad r(R, Z) = -0,485 \quad r(X, Z) = 0,993. \quad (2.58)$$

2/ En calculant d'abord une valeur pour R, X et Z pour chaque ensemble de données d'entrée et en les utilisant ensuite pour obtenir les meilleures estimations des grandeurs demandées. Les résultats sont identiques.

2.5 Niveaux de confiance et degrés de liberté

2.5.1 Problématique

Il est parfois nécessaire de s'assurer qu'un grand nombre des valeurs obtenues (95% ou 99% par exemple) soient bien comprises dans un intervalle donné. Dans le cas idéal, il suffit de multiplier l'écart-type par un **facteur d'élargissement** [coverage factor] pour obtenir une **incertitude élargie** [expanded uncertainty] correspondant à un **niveau de confiance spécifié** [specified level of confidence]. Voir figure 2.2. Le tableau 2.3 donne la correspondance entre le niveau de confiance et le facteur d'élargissement pour une loi normale.

Mais, comment passer de la variance à un intervalle de confiance quand on est pas sûr que la loi de probabilité soit normale et qu'il y a une incertitude sur la variance ? Nous avons vu qu'avec 30 mesures, il y a encore une incertitude relative de 13% sur un écart-type. Même si l'on est sûr d'avoir une loi de probabilité normale, on

Niveau de confiance p (pourcentage)	Facteur d'élargissement k_p
68,27	1
90	1,645
95	1,960
95,45	2
99	2,576
99,73	3

TAB. 2.3 – Valeur du facteur d'élargissement k_p qui donne un intervalle correspondant au niveau de confiance p , en supposant la loi normale.

ne distinguera pas 94, 95 ou 96%. Si le théorème de la limite centrale permet d'affirmer que l'on tend vers une loi normale pour un grand nombre de degrés de libertés (plus de 20 généralement), ces conditions ne sont pas toujours remplies et il faut définir d'autres outils statistiques.

2.5.2 Table ou loi de Student

Soit z une variable aléatoire d'espérance mathématique μ_z et d'écart-type σ suivant une loi normale. Si \bar{z} est la moyenne arithmétique de n observations indépendantes z_k de z et $s(\bar{z})$, l'écart-type expérimental, alors la variable t ,

$$t = \frac{\bar{z} - \mu_z}{s(\bar{z})}, \quad (2.59)$$

suit la **loi de t** ou **loi de Student** :

$$p(t, \nu) = \frac{1}{\sqrt{\pi\nu}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left[1 + \frac{t^2}{\nu}\right]^{-(\nu+1)/2}, \quad (2.60)$$

où Γ est la fonction Gamma et $\nu = n - 1$ est le nombre de degrés de liberté. Comme cette expression est difficile à manipuler, le GUM propose une table, dite **table de Student**, qui donne en fonction de ν et p les valeurs $t_p(\nu)$ telles que

$$Pr[-t_p(\nu) \leq t \leq t_p(\nu)] = p. \quad (2.61)$$

Ici, $Pr[\]$ signifie "probabilité de" et l'équation "la probabilité que $t \in [-t_p(\nu), t_p(\nu)]$ est égale à p ". En remplaçant t par son expression, on obtient

$$Pr[\bar{z} - t_p(\nu)s(\bar{z}) \leq \mu_z \leq \bar{z} + t_p(\nu)s(\bar{z})] = p. \quad (2.62)$$

La table de Student donne accès au facteur d'élargissement pour une probabilité et un nombre de degrés de liberté donnés. Une sélection de valeurs est donnée dans le tableau 2.4. Une table plus complète est disponible dans la norme.

Pour rappel, la fonction Γ est aussi appelée *intégrale d'Euler de deuxième espèce* et est définie par

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt. \quad (2.63)$$

On a facilement que $\Gamma(1) = 1$ et en intégrant par partie que $\Gamma(x+1) = x\Gamma(x)$, $\forall x > 0$. Ainsi, pour n entier, on a $\Gamma(n+1) = n!$. Comme la formule de Student fait intervenir des demi-entiers, $\Gamma(1/2) = \sqrt{\pi}$ peut être utile. Le calcul se fait avec le changement de variable $u = \sqrt{t}$.

2.5.3 Nombre de degrés de liberté

Le **nombre de degrés de liberté** [the degrees of freedom] devient un élément clé de l'estimation de l'incertitude. Il est égal à $n - 1$ pour une grandeur unique estimée par la moyenne arithmétique de n observations

ν	$p = 68, 27\%$	$p = 95\%$	$p = 99\%$	$p = 99, 73\%$
1	1,84	12,71	63,66	235,80
2	1,32	4,30	9,92	19,21
3	1,20	3,18	5,84	9,22
4	1,14	2,78	4,60	6,62
5	1,11	2,57	4,03	5,51
10	1,05	2,23	3,17	3,96
20	1,03	2,09	2,85	3,42
50	1,01	2,01	2,68	3,16
100	1,005	1,984	2,626	3,077
∞	1,000	1,960	2,576	3,000

TAB. 2.4 – Valeur du facteur d’élargissement $t_p(\nu)$ en fonction du nombre de degrés de liberté ν .

indépendantes. Si l’on détermine à la fois une pente et une ordonnée à l’origine par la méthode des moindres carrés, $\nu = n - 2$. Si m paramètres sont ajustés sur n valeurs par la méthode des moindres carrés, $\nu = n - m$.

Pour une incertitude de type B, on définit le nombre de degrés de liberté par

$$\nu_i \simeq \frac{1}{2} \left[\frac{\Delta u(x_i)}{u(x_i)} \right]^{-2}, \quad (2.64)$$

où $\Delta u(x_i)/u(x_i)$ est une incertitude relative de $u(x_i)$. C’est une grandeur subjective dont la valeur s’obtient par “jugement d’expert” fondé sur l’ensemble des informations disponibles. Par exemple, si une incertitude-type est jugée fiable à 10% cela conduit à un nombre de degrés de liberté estimé $\nu_i = 50$. Si la loi de probabilité est rectangulaire, les bornes sont généralement choisies de façon à inclure tout l’intervalle de probabilité et il n’est donc pas irréaliste de prendre $\nu_i \rightarrow \infty$.

Le problème se pose quand on traite une grandeur y dépendant de N grandeurs d’entrée x_i normalement distribuées, $y = f(x_1, \dots, x_N)$. On peut faire l’approximation que la loi de Student s’applique avec un nombre de degrés de liberté effectif ν_{eff} donné par la loi de Welch-Satterthwaite [12, 13],

$$\frac{\sigma^4(y)}{\nu_{eff}} = \sum_{i=1}^N \frac{\left[\frac{\partial f}{\partial x_i} \sigma(x_i) \right]^4}{\nu_i}. \quad (2.65)$$

Attention, cette loi est exprimée différemment dans la norme en introduisant $u_i(y) = \frac{\partial f}{\partial x_i} \sigma(x_i)$:

$$\frac{u_c^4(y)}{\nu_{eff}} = \sum_{i=1}^N \frac{u_i^4(y)}{\nu_i}. \quad (2.66)$$

Dans la pratique, on arrondit généralement ν_{eff} à l’entier inférieur. On a $\nu_{eff} \leq \sum_{i=1}^N \nu_i$. Si une des grandeurs x_i est elle aussi estimée à partir d’autres grandeurs, on utilisera un nombre de degrés de liberté effectif pour elle, estimé de façon analogue.

On peut alors utiliser la table de Student pour trouver le facteur d’élargissement $t_p(\nu_{eff})$ à appliquer pour une grandeur de sortie obtenue à partir d’autres grandeurs d’entrée.

La formule approximative de Welch-Satterthwaite a été dérivée pour des incertitudes de type A et a été extrapolée par le GUM aux incertitudes de type B. Il existe des articles scientifiques montrant que dans certains cas, elle ne peut pas être utilisée.

Exemple tiré de la norme

Supposons que $Y = bX_1X_2X_3$ et que les valeurs d’entrée x_1, x_2, x_3 , normalement distribuées, sont estimées par leur moyennes arithmétiques de $n_1 = 10$, $n_2 = 5$ et $n_3 = 15$ observations répétées indépendantes. Les incertitudes-types relatives sont $s(x_1)/x_1 = 0,25\%$, $s(x_2)/x_2 = 0,57\%$ et $s(x_3)/x_3 = 0,82\%$. On a

immédiatement pour y ,

$$\left[\frac{u_c(y)}{y} \right]^2 = \left[\frac{s(x_1)}{x_1} \right]^2 + \left[\frac{s(x_2)}{x_2} \right]^2 + \left[\frac{s(x_3)}{x_3} \right]^2 = (1,03\%)^2. \quad (2.67)$$

De la même façon, du fait de la forme de la fonction f , on peut directement utiliser les incertitudes-types relatives pour calculer ν_{eff} , mais ce n'est pas toujours possible. Cet exemple peut donc être trompeur.

$$\nu_{eff} = \frac{(u_c(y)/y)^4}{\frac{(s(x_1)/x_1)^4}{\nu_1} + \frac{(s(x_2)/x_2)^4}{\nu_2} + \frac{(s(x_3)/x_3)^4}{\nu_3}} = \frac{1,03^4}{\frac{0,25^4}{10-1} + \frac{0,57^4}{5-1} + \frac{0,82^4}{15-1}} = 19,0. \quad (2.68)$$

D'après la table Student, $t_{95\%}(19) = 2,09$. L'incertitude élargie relative pour un niveau de confiance de 95% est $U_{95} = 2,09 \times 1,03\% = 2,2\%$.

2.6 Recommandations pour la présentation des résultats

Les résultats exprimés doivent pouvoir être utilisés facilement par le lecteur sans qu'il ait à refaire des calculs difficiles. Cela signifie qu'il ne faut pas avoir peur d'être redondant. Je vais reprendre ici la recommandation internationale INC-1 (1980) **Expression des incertitudes expérimentales** :

1. L'incertitude d'un résultat de mesure comprend généralement plusieurs composantes qui peuvent être généralement groupées en deux catégories d'après la méthode utilisée pour estimer leur valeur numérique :

- A** celles qui sont évaluées à l'aide de méthodes statistiques,
- B** celles qui sont évaluées par d'autres moyens.

Il n'y a pas toujours de correspondance simple entre le classement dans les catégories A ou B et le caractère "aléatoire" ou "systématique" utilisé antérieurement pour classer les incertitudes. L'expression "incertitude systématique" est susceptible de conduire à des erreurs d'interprétation : elle doit être évitée. Toute description détaillée de l'incertitude devrait comprendre une liste complète de ses composantes et indiquer pour chacune la méthode utilisée pour lui attribuer une valeur numérique.

2. Les composantes de la catégorie A sont caractérisées par des variances estimées s_i^2 (ou les "écarts-types" estimés s_i) et les nombres ν_i de degré de liberté. Le cas échéant, les covariances estimées doivent être données.
3. Les composantes de la catégorie B devraient être caractérisées par les variances estimées u_j^2 , qui puissent être considérées comme des approximations des variances correspondantes dont on admet l'existence. Les termes u_j^2 peuvent être traités comme des variances et les termes u_j comme des écarts-types. Le cas échéant, les covariances doivent être traitées de façon analogue.
4. L'incertitude composée devrait être caractérisée par la valeur obtenue en appliquant la méthode usuelle de combinaison des variances. L'incertitude composée ainsi que ses composantes devraient être exprimées sous la forme d'"écarts-types".
5. Si, pour des utilisations particulières, on est amené à multiplier par un facteur l'incertitude composée afin d'obtenir une incertitude globale, la valeur numérique de ce facteur doit toujours être donnée.

2.7 Résumé et conclusion

La procédure proposée par le GUM repose sur 4 étapes¹ :

1. **Ecrire le modèle mathématique du processus de mesure.** Après avoir analysé le processus de mesure, identifié les facteurs d'influence et établi une stratégie pour les réduire (répétition des observations et application de corrections). La transcription de ces opérations sous forme mathématique conduit à l'établissement du modèle mathématique.

¹Ce résumé est tiré du livre de C. Perruchet et M. Priel.

2. **Le calcul des incertitudes-types.** Pour chacune des grandeurs d'entrée du modèle, il existe un doute (une incertitude). Ces incertitudes sont évaluées, en utilisant des méthodes de type A, estimation d'un écart-type à partir d'une série d'observations, ou des méthodes de type B, estimation d'un écart-type se fondant sur une évaluation de l'étendue de variation possible pour la grandeur d'entrée considérée et choix d'une forme de distribution.
3. **La propagation des incertitudes.** Chacune des grandeurs d'entrée contribue à l'incertitude sur le résultat. La loi de propagation des variances (ou des incertitudes) permet de calculer la variance sur le résultat.
4. **L'expression du résultat et de son incertitude.** Connaissant la variance sur le résultat, il est possible d'exprimer l'incertitude.

En guise de conclusion, voici le paragraphe 3.4.8 de la norme NF ENV 13005 :

Bien que ce *Guide* fournisse un cadre pour l'estimation de l'incertitude, il ne peut remplacer ni la réflexion critique ni l'honnêteté intellectuelle ni la compétence professionnelle. L'évaluation de l'incertitude n'est jamais une tâche de routine ni une opération purement mathématique ; elle dépend de la connaissance détaillée de la nature du mesurande et du mesurage. La qualité et l'utilité de l'incertitude fournie pour le résultat d'un mesurage dépend, en fin de compte, de la compréhension, de l'analyse critique et de l'intégrité de ceux qui contribuent à son évaluation.

Chapitre 3

Calcul de la propagation de l'incertitude en utilisant des méthodes Monte Carlo

La formule de propagation des incertitudes, équation (2.24), est basée sur un développement de Taylor. Si les incertitudes types relatives des grandeurs d'entrée ne sont pas très petites ou si la relation entre les grandeurs d'entrée et la grandeur de sortie présente de fortes non-linéarités, ce développement peut ne pas être valable. L'estimation du facteur d'élargissement est basé sur une hypothèse gaussienne ou sur la loi de Student. Mais si une incertitude d'entrée a une distribution rectangulaire et domine le calcul d'incertitudes ou si les incertitudes de type A sont évaluées avec un très petit nombre de données, il est alors difficile de lier un niveau de confiance et un facteur d'élargissement. Pour ces cas en particulier, mais aussi pour valider une formule analytique, on peut faire une estimation numérique de l'incertitude. La démarche à suivre est décrite dans :

Evaluation des données de mesure - Supplément 1 du "Guide pour l'expression de l'incertitude de mesure" - Propagation de distributions par une méthode de Monte Carlo (2008). Une version en anglais est disponible gratuitement sur le site du BIPM : <http://www.bipm.org/>.

La résolution numérique est indispensable quand les grandeurs d'entrée et la grandeur de sortie sont liées par un code numérique et pas une équation.

3.1 Démarche

3.1.1 Formulation du problème

Supposons que la grandeur de sortie Y dépende de N grandeurs d'entrée notées $[X] = (X_1, \dots, X_N)^T$. Y est un scalaire, alors que $[X]$ est un vecteur vertical (d'où la transposition notée T). La relation entre Y et $[X]$ se fait généralement à l'aide d'une fonction f ,

$$Y = f([X]). \quad (3.1)$$

mais elle peut aussi être exprimée de façon implicite, $h(Y, [X]) = 0$. Dans ce chapitre, les grandeurs entre crochets sont des vecteurs ou des matrices. Elles sont en caractères gras dans le supplément au GUM.

Chaque grandeur d'entrée X_i est considérée comme une **variable aléatoire** [random variable] pouvant prendre les valeurs ξ_i et ayant x_i comme valeur moyenne.

Pour une seule grandeur d'entrée X , la **fonction densité de probabilité** [probability density function, PDF] est notée $g_X(\xi)$: $g_X(\xi) d\xi$ représente la probabilité que $X \in [\xi, \xi + d\xi]$. Ainsi, l'espérance et la variance valent

$$E(X) = \int_{-\infty}^{+\infty} \xi g_X(\xi) d\xi \quad (3.2)$$

$$V(X) = \int_{-\infty}^{+\infty} (\xi - E(X))^2 g_X(\xi) d\xi. \quad (3.3)$$

S'il y a deux grandeurs d'entrée, X_1 et X_2 , a priori corrélées, on n'a qu'une seule PDF pour $[X]$ notée $g_{[X]}([\xi])$. Avec une loi normale, elle s'écrit

$$g_{[X]}([\xi]) = \frac{1}{2\pi} \frac{1}{\sqrt{\sigma_1\sigma_2 - \sigma_{12}^2}} \exp\left(-\frac{\sigma_2(\xi_1 - x_1)^2 + \sigma_1(\xi_2 - x_2)^2 - 2\sigma_{12}(\xi_1 - x_1)(\xi_2 - x_2)}{2(\sigma_1\sigma_2 - \sigma_{12}^2)}\right) \quad (3.4)$$

$$= \frac{1}{2\pi} \frac{1}{\sqrt{\det[\sigma]}} \exp\left(-\frac{1}{2}[\xi - x]^T \cdot [\sigma]^{-1} \cdot [\xi - x]\right). \quad (3.5)$$

Si les deux grandeurs d'entrée sont indépendantes, alors

$$g_{[X]}([\xi]) = g_{X_1}(\xi_1) \cdot g_{X_2}(\xi_2). \quad (3.6)$$

Mais si elles sont corrélées, on ne peut pas définir une PDF pour chaque variable. La covariance entre X_1 et X_2 vaut

$$Cov(X_1, X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\xi_1 - E(X_1))(\xi_2 - E(X_2)) g_{[X]}([\xi]) d\xi_1 d\xi_2. \quad (3.7)$$

L'espérance de $[X]$ est aussi un vecteur $E([X])$ de coordonnées

$$E(X_i) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \xi_i g_{[X]}([\xi]) d\xi_1 d\xi_2, \quad (3.8)$$

et la variance devient une matrice 2×2 ,

$$V([X]) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [X - E([X])] \cdot [X - E([X])]^T g_{[X]}([\xi]) d\xi_1 d\xi_2. \quad (3.9)$$

A partir de deux grandeurs d'entrée, on peut généraliser aisément à N grandeurs.

La grandeur de sortie, Y sera aussi une variable aléatoire pouvant prendre les valeurs η et ayant une valeur moyenne notée y .

3.1.2 Les étapes de l'évaluation

Formulation du problème

On définit le mesurande (Y) et les grandeurs d'entrée ($[X]$) ainsi qu'un modèle reliant Y et $[X]$. A chaque grandeur d'entrée, on associe une distribution (normale, rectangulaire, ...) si elles sont indépendantes, ou une distribution conjointe, si elles sont corrélées.

Propagation

On effectue M tirages aléatoires pour chacune des grandeurs d'entrée et on en déduit M valeurs pour Y . M est typiquement de l'ordre de 10^6 .

Synthèse

A partir de la PDF pour Y , on peut calculer l'espérance mathématique, l'écart-type et l'incertitude élargie correspondant au niveau de confiance demandé. Toutes ces données doivent apparaître dans l'expression des résultats.

Exemple :

$$y = 1,024 \text{ V}, \quad u(y) = 0,028 \text{ V}$$

Intervalle le plus court pour un niveau de confiance de 95% = [0,983, 1,088].

3.2 Choix et construction des distributions d'entrée

Quand on ne connaît pas la PDF d'une grandeur d'entrée, mais seulement quelques informations, il faut faire un choix parmi toutes celles qui satisfont aux données. Et ce choix ne doit pas dépendre de la personne qui le fait. Le supplément à la norme propose de nombreux exemples obtenus à partir du principe de maximisation de l'entropie. Voir chapitre 6 pour en savoir plus sur ce principe et son application.

Je ne vais reprendre ici que quelques exemples simples à titre d'illustration. Je vous renvoie au supplément 1 du GUM pour les autres possibilités.

3.2.1 Distribution rectangulaire

Si la seule information concernant une grandeur X est la limite inférieure a et la limite supérieure b (avec $b > a$), alors une distribution rectangulaire doit être affectée à X sur l'intervalle $[a, b]$.

La PDF est alors,

$$g_X(\xi) = \begin{cases} 1/(b-a), & \text{si } \xi \in [a, b] \\ 0, & \text{autrement} \end{cases} . \quad (3.10)$$

L'espérance mathématique et la variance valent alors,

$$E(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a)^2}{12}. \quad (3.11)$$

Pour avoir une distribution rectangulaire $R(a, b)$ à partir de r tiré suivant une distribution rectangulaire $R(0, 1)$, c'est très simple,

$$\xi = a + (b-a)r. \quad (3.12)$$

3.2.2 Distribution trapézoïdale

Si la grandeur X est la somme de deux grandeurs indépendantes X_1 et X_2 ayant chacune une distribution rectangulaire, alors la distribution pour X sera trapézoïdale. Les formules équivalentes à l'exemple précédent sont dans le supplément au GUM.

3.2.3 Distribution gaussienne simple

Si la seule information disponible est la valeur moyenne x et l'écart-type standard $u(x)$, alors la distribution doit être choisie gaussienne. La démonstration est faite dans le chapitre 6.

La PDF pour X s'écrit,

$$g_X(\xi) = \frac{1}{\sqrt{2\pi}u(x)} \exp\left(-\frac{(\xi-x)^2}{2u^2(x)}\right). \quad (3.13)$$

L'espérance mathématique et la variance sont, bien-entendu, $E(X) = x$ et $V(X) = u^2(x)$.

Numériquement, ξ peut être obtenu à partir d'un tirage z selon une distribution normale de valeur moyenne nulle et de variance égale à 1, de façon très simple,

$$\xi = x + u(x)z. \quad (3.14)$$

Cette distribution gaussienne sur z peut être obtenue à partir d'un tirage rectangulaire et une transformation de Box-Muller. A partir de r_1 et r_2 tirés suivant une distribution rectangulaire $R(0, 1)$, on obtient deux nombres

$$z_1 = \sqrt{-2 \ln r_1} \cos(2\pi r_2) \quad \text{et} \quad z_2 = \sqrt{-2 \ln r_2} \cos(2\pi r_1), \quad (3.15)$$

ayant une distribution gaussienne.

3.2.4 Distribution gaussienne à plusieurs variables

Méthode

Pour un ensemble de N grandeurs d'entrée $[X] = (X_1, \dots, X_N)^T$ ayant une valeur moyenne $[x] = (x_1, \dots, x_N)^T$ et une matrice d'incertitude associée

$$[U] = \begin{bmatrix} u^2(x_1) & u(x_1, x_2) & \cdots & u(x_1, x_N) \\ u(x_2, x_1) & u^2(x_2) & \cdots & u(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ u(x_N, x_1) & u(x_N, x_2) & \cdots & u^2(x_N) \end{bmatrix}, \quad (3.16)$$

la PDF s'écrit,

$$g_{[X]}([\xi]) = \frac{1}{\sqrt{(2\pi)^N \det[U]}} \exp\left(-\frac{1}{2}[\xi - x]^T \cdot [U]^{-1} \cdot [\xi - x]\right). \quad (3.17)$$

Evidemment, l'espérance est $E([X]) = [x]$ et la variance $V(X) = [U]$.

Pour décorréler les N grandeurs d'entrée, on peut diagonaliser la matrice $[U]$, mais si elle est bien symétrique définie positive, il est plus simple de faire une factorisation de Cholesky, nommée d'après André-Louis Cholesky (1875 - 1918), qui consiste à déterminer une matrice triangulaire supérieure $[R]$ telle que $[U] = [R]^T \cdot [R]$. Cela se fait par récurrence et est très facile à numériser. Finalement, on tire N nombres z_i indépendants suivant une distribution normale de valeur moyenne nulle et de variance égale à 1 pour obtenir

$$[\xi] = [x] + [R]^T \cdot [z]. \quad (3.18)$$

Factorisation de Cholesky

Prenons un exemple, ce sera plus simple qu'avec des formules générales. On veut résoudre,

$$\begin{pmatrix} 1 & 2 & -3 \\ 2 & 13 & 6 \\ -3 & 6 & 50 \end{pmatrix} = \begin{pmatrix} r_{11} & 0 & 0 \\ r_{12} & r_{22} & 0 \\ r_{13} & r_{23} & r_{33} \end{pmatrix} \cdot \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}. \quad (3.19)$$

En effectuant le produit des deux matrices de gauche, on arrive au système suivant,

$$\begin{aligned} 1 &= r_{11}^2 && \Rightarrow r_{11} = 1 \\ 2 &= r_{11}r_{12} && \Rightarrow r_{12} = 2 \\ -3 &= r_{11}r_{13} && \Rightarrow r_{13} = -3 \\ 13 &= r_{12}^2 + r_{22}^2 && \Rightarrow r_{22} = 3 \\ 6 &= r_{12}r_{13} + r_{22}r_{23} && \Rightarrow r_{23} = 4 \\ 50 &= r_{13}^2 + r_{23}^2 + r_{33}^2 && \Rightarrow r_{33} = 5 \end{aligned} \quad (3.20)$$

qui peut être résolu ligne par ligne, comme indiqué. En conclusion,

$$[R] = \begin{pmatrix} 1 & 2 & -3 \\ 0 & 3 & 4 \\ 0 & 0 & 5 \end{pmatrix}. \quad (3.21)$$

Même pour des matrices plus grandes, la procédure est facile à numériser. Il existe des sous-routines toutes faites dans les bibliothèques.

3.2.5 Générateur de nombres pseudo-aléatoires

Il est important d'avoir un générateur performant de nombres pseudo-aléatoires ayant une distribution rectangulaire. La qualité des autres distributions en dépend.

Un bon générateur doit avoir les bonnes propriétés statistiques, pouvoir être implémenté dans n'importe quel langage de programmation et donner les mêmes résultats avec tout ordinateur. Sauf si l'utilisateur est sûr de son pedigree, un tel générateur ne doit pas être utilisé sans quelques tests de performance.

Le supplément suggère d'utiliser le générateur de nombres pseudo-aléatoires amélioré de Wichmann-Hill [enhanced Wichmann-Hill generator for producing pseudo-random numbers]. Voir le supplément 1 du GUM pour plus de détails.

3.3 Mise en œuvre pratique

3.3.1 Tolérance numérique

Soit n_{dig} le nombre de chiffres significatifs considérés comme pertinents lors d'une évaluation numérique de z . On écrit z de la forme $z = c \times 10^l$. Ici, c est un entier de n_{dig} chiffres et l est un entier.

La **tolérance numérique** [numerical tolerance] est définie par $\delta = \frac{1}{2}10^l$.

Exemple : Pour une masse $m = 100,021\,47 \pm 0,000\,35$ g, le nombre de chiffre significatifs est 2 pour l'incertitude. Ici $c = 35$ et $l = -5$. La tolérance numérique est donc $\delta = \frac{1}{2} \times 10^{-5}$ g = 0,000 005 g.

3.3.2 Stabilité numérique

Une évaluation de type Monte Carlo de l'incertitude nécessite d'effectuer un grand nombre de calculs test jusqu'à ce que les grandeurs intéressantes soient stables numériquement. Un résultat numérique est reconnu comme stable numériquement si deux fois l'écart-type lié aux fluctuations numériques est inférieur à la tolérance numérique de l'incertitude type $u(y)$.

3.3.3 Algorithme

L'objectif de la procédure dite adaptative, présentée ci-dessous, est de donner une estimation y de Y , l'incertitude type associée $u(y)$ et les extrémités de l'intervalle élargi y^{\triangleleft} et y^{\triangleright} associé à une probabilité de couverture p . La stabilité numérique est recherchée pour chacune de ces quatre valeurs.

A noter que y et $u(y)$ convergent beaucoup plus vite que y^{\triangleleft} et y^{\triangleright} quand on augmente M . Une probabilité de couverture plus grande nécessitera un plus grand nombre de tests Monte Carlo.

Voici la procédure adaptative proposée par le supplément au GUM :

1. Fixer n_{dig} ;
2. Choisir $M = \max(\frac{100}{1-p}, 10^4)$;
3. Initialiser h à $h = 1$;
4. Calculer M valeurs de Y , pour calculer $y_h, u(y_h), y_h^{\triangleleft}$ et y_h^{\triangleright} ;
5. Si $h = 1$, augmenter h de 1 et retourner à l'étape 4 ; sinon, passer à l'étape suivante ;
6. Calculer la variance associée aux estimations y_1, \dots, y_h de Y ,

$$s_y^2 = \frac{1}{h(h-1)} \sum_{r=1}^h (y_r - y)^2 \quad \text{avec} \quad y = \frac{1}{h} \sum_{r=1}^h y_r ; \quad (3.22)$$

7. Faire la même chose pour $u(y), y^{\triangleleft}$ et y^{\triangleright} ;
8. Utiliser les $h \times M$ valeurs disponibles pour calculer $u(y)$ et calculer la tolérance numérique associée δ ;
9. Si l'une des quatre valeurs, $2s_y, 2s_{u(y)}, 2s_{y^{\triangleleft}}$ ou $2s_{y^{\triangleright}}$, dépasse δ , augmenter h de 1 et retourner à l'étape 4 ; sinon, passer à l'étape suivante ;
10. Considérer que l'évaluation Monte Carlo s'est stabilisée et utiliser les $h \times M$ valeurs disponibles pour calculer $y, u(y), y^{\triangleleft}$ et y^{\triangleright} .

Quelques remarques concernant cette procédure :

- Le choix de M est arbitraire et vient de la pratique.
- Les écarts-types calculés aux étapes 6 et 7 diminuent en $1/\sqrt{h}$.
- Si l'intervalle élargi n'est pas demandé, on peut enlever le test correspondant et cela convergera plus vite.
- Le facteur d'élargissement de 2 choisi à l'étape 9 correspond à une probabilité de couverture de 95%. Pour une telle probabilité, généralement $M_{tot} = h \times M \simeq 10^6$.

3.4 Validation du GUM

Nous l'avons dit en introduction, les conditions d'application de la formule de propagation des incertitudes du GUM sont limitées. La méthode Monte Carlo peut être utilisée pour valider ou non un calcul analytique type GUM qui est plus simple.

En notant U_p l'incertitude élargie obtenue par la formule du GUM, si $|y - U_p - y^\triangleleft|$ et $|y + U_p - y^\triangleright|$ sont tous les deux inférieurs à la tolérance numérique, alors, on peut considérer l'approche du GUM comme validée. Il est important de noter que pour faire ce test, le code Monte Carlo doit tourner jusqu'à avoir une stabilité numérique estimée avec une tolérance numérique égale à $\delta/5$. Cela signifie une statistique 25 fois plus élevée.

Si la comparaison est satisfaisante, la méthode du GUM pourra être utilisée à cette occasion ou pour des problèmes similaires.

Chapitre 4

Traitement statistique des données

Pour tirer des informations de données expérimentales on les compare souvent à des modèles théoriques ou bien, on les utilise pour ajuster les paramètres d'un modèle phénoménologique. Il existe des méthodes statistiques d'ajustement [fit] qui font l'objet de ce chapitre. Une fois les paramètres ajustés, il faut évaluer leur incertitude et la qualité de l'ajustement.

Ce chapitre est basé essentiellement sur l'excellent livre de John R. Taylor, *An Introduction to Error Analysis : The Study of Uncertainties in Physical Measurements* (University Science Books, 1997) [6] dont il existe une version française.

Les prémices des méthodes d'ajustement sont basées sur des notions élémentaires de probabilité que je vais commencer par présenter.

4.1 Préambule : quelques notions sur les probabilités

La probabilité de tirer un as d'un jeu de 52 cartes est de $p_A = 4/52$. La probabilité de tirer une carte rouge de ce même jeu est de $p_R = 1/2$. La probabilité de tirer un as rouge d'un jeu de 52 cartes est de $2/52 = p_A \times p_R$.

D'une manière générale, si l'on note $p(A)$ la probabilité que la proposition A soit vraie et $p(B)$ celle que la proposition B soit vraie, on note alors AB ou $A \cap B$ la proposition que A ET B soient vraies. Avec l'exemple précédent, A pourrait être « la carte tirée est un As » et la B « la carte tirée est rouge ». La figure 4.1 montre une représentation schématique de la condition ET. En bleu, on a la zone où A est vraie et en jaune, celle où B est vraie. La zone où A et B sont vraies correspond à l'intersection représentée en vert. On a

$$p(AB) = p(A \cap B) = p(A) \cdot p(B). \quad (4.1)$$

Cette relation n'est vraie que si A et B sont indépendantes. Nous verrons au chapitre 6 le cas où A et B sont corrélées.

Passons à « pile ou face ». D'après ce que nous venons de voir, la probabilité que le tirage 1 donne face Et que le tirage 2 donne aussi face est $p(1) \cdot p(2) = 1/4$. Qu'en est-il de la proposition « le tirage 1 OU le tirage 2 » donne face ? Elle est notée $A + B$ ou $A \cup B$. La règle d'addition s'écrit

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = p(A) + p(B) - p(A) \cdot p(B), \quad (4.2)$$

car il ne faut pas compter deux fois l'intersection (en vert sur la figure 4.1). Ainsi, la probabilité de tirer au moins une face sur deux tirages est de $3/4$.

Ces règles ont été illustrées par des exemples simples ayant une interprétation fréquentiste. Elles restent valable dans les autres cas, comme « la probabilité qu'il y ait de la vie sur Mars », où cette interprétation n'est plus possible. R.T. Cox a démontré, dans un article célèbre [14], que ces règles peuvent être déduites d'une axiomatique très simple. Voir aussi l'article pédagogique de W. Tyler Estler [8] et les références qu'il contient.

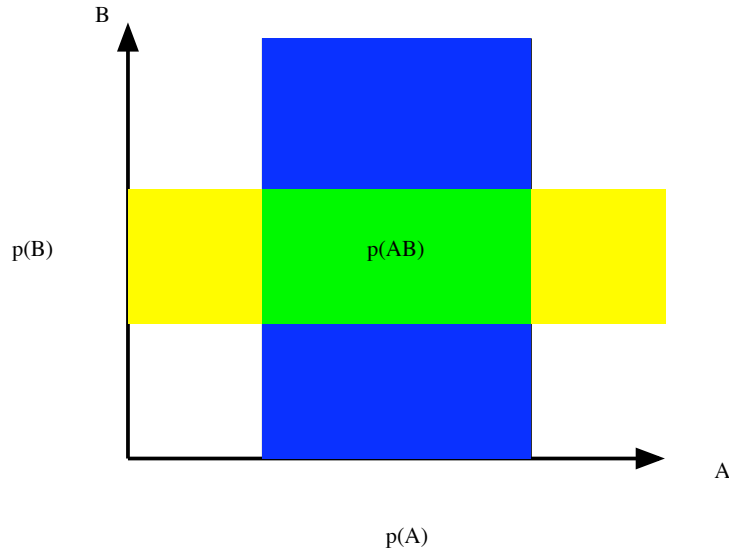


FIG. 4.1 – Représentation de la probabilité que A ET B soient vraies.

4.2 Moyenne pondérée

Imaginons que le même mesurande ait fait l'objet de deux mesurages indépendants dans deux laboratoires différents par exemple. Les résultats obtenus sont

$$x = x_A \pm \sigma_A \quad \text{dans le labo A et} \quad x = x_B \pm \sigma_B \quad (4.3)$$

dans le labo B. Comment combiner ces résultats pour avoir une estimation encore meilleure ?

Si $|x_A - x_B|$ est beaucoup plus grand que σ_A et σ_B , alors ces deux résultats sont inconsistants. Supposons qu'ils soient consistants.

La probabilité que la labo A mesure la valeur x est égale à

$$P(x|x_A, \sigma_A) = \frac{1}{\sqrt{2\pi}\sigma_A} \exp\left(-\frac{(x-x_A)^2}{2\sigma_A^2}\right). \quad (4.4)$$

De même pour le labo B. La probabilité que la labo A mesure x ET que le labo B mesure aussi x est donc le produit de ces deux probabilités, à savoir

$$P(x|x_A, \sigma_A, x_B, \sigma_B) = \frac{1}{2\pi\sigma_A\sigma_B} \exp\left(-\frac{\chi^2}{2}\right) \quad \text{avec} \quad \chi^2 = \frac{(x-x_A)^2}{\sigma_A^2} + \frac{(x-x_B)^2}{\sigma_B^2}. \quad (4.5)$$

La meilleure estimation de X correspond donc au maximum de cette probabilité, c'est à dire à la valeur qui minimise χ^2 . En dérivant par rapport à x , on obtient finalement que la meilleure estimation de X vaut,

$$x_m = \frac{\frac{x_A}{\sigma_A^2} + \frac{x_B}{\sigma_B^2}}{\frac{1}{\sigma_A^2} + \frac{1}{\sigma_B^2}}. \quad (4.6)$$

Il s'agit de la moyenne pondérée des coefficients $w_A = 1/\sigma_A^2$ et $w_B = 1/\sigma_B^2$ respectivement,

$$x_m = \frac{w_A x_A + w_B x_B}{w_A + w_B}. \quad (4.7)$$

On peut généraliser aisément à N mesurages de X :

$$x_m = \frac{\sum_{k=1}^N w_k x_k}{\sum_{k=1}^N w_k} \quad \text{avec} \quad w_k = \frac{1}{\sigma_k^2}. \quad (4.8)$$

Ainsi, un résultat moins précis aura un poids plus faible dans la moyenne.

L'incertitude correspondant à cette nouvelle valeur peut être calculée facilement à l'aide de la formule de propagation des incertitudes du GUM. On trouve, pour la variance,

$$\sigma_{x_m}^2 = \frac{1}{\sum_{k=1}^N w_k}. \quad (4.9)$$

4.3 Méthode des moindres carrés

Supposons maintenant que l'on a une série de données expérimentales que l'on veut "fitter" avec une courbe. Pour un type de courbe donné (droite, polynôme, . . .), quelle est la courbe la plus probable? Quelle est l'incertitude sur le résultat (coefficients de la fonction)?

Nous allons d'abord traiter le cas le plus simple, une droite, puis nous verrons des cas plus compliqués.

La **méthode des moindres carrés** [least squares fitting] a été élaborée indépendamment par Adrien-Marie Legendre en 1805 et Carl Friedrich Gauss en 1809.

4.3.1 Régression linéaire

Considérons un ensemble de n données (x_i, y_i) qui devraient s'aligner sur une droite d'équation $y = Ax + B$. On cherche à déterminer les coefficients A et B les plus probables et les incertitudes associées.

Supposons, pour commencer par un cas simple, que l'on peut négliger les incertitudes sur les x_i et que les y_i ont une incertitude $u(y_i)$. On connaît alors la valeur vraie de chaque y_i qui est $Ax_i + B$.

La probabilité que les coefficients A et B conviennent pour un couple (x_i, y_i) est donc

$$P(A, B|x_i, y_i, u(y_i)) = \frac{1}{\sqrt{2\pi}u(y_i)} \exp\left(-\frac{(y_i - Ax_i - B)^2}{2u^2(y_i)}\right). \quad (4.10)$$

Et donc, la probabilité que A et B conviennent pour tous les couples est

$$P(A, B|x_1, \dots, x_n, y_1, \dots, y_n, u(y_1), \dots, u(y_n)) = \frac{1}{\sqrt{(2\pi)^n \prod_{k=1}^n u(y_k)}} \exp\left(-\frac{\chi^2}{2}\right). \quad (4.11)$$

avec

$$\chi^2 = \sum_{k=1}^n \frac{(y_k - Ax_k - B)^2}{u^2(y_k)} = \sum_{k=1}^n w_k (y_k - Ax_k - B)^2 \quad (4.12)$$

en introduisant les facteurs de pondération $w_k = 1/u^2(y_k)$.

A et B sont déterminés en minimisant χ^2 , d'où le nom de méthode des moindres carrés [least-squares fitting]. En dérivant χ^2 par A et B respectivement, on tombe sur un système de deux équations à deux inconnues,

$$A \sum_{k=1}^n w_k x_k^2 + B \sum_{k=1}^n w_k x_k - \sum_{k=1}^n w_k x_k y_k = 0 \quad (4.13)$$

$$A \sum_{k=1}^n w_k x_k + B \sum_{k=1}^n w_k - \sum_{k=1}^n w_k y_k = 0 \quad (4.14)$$

facile à résoudre. On trouve,

$$A = \frac{1}{\Delta} \left(\sum_{k=1}^n w_k \cdot \sum_{k=1}^n w_k x_k y_k - \sum_{k=1}^n w_k x_k \cdot \sum_{k=1}^n w_k y_k \right) \quad (4.15)$$

$$B = \frac{1}{\Delta} \left(\sum_{k=1}^n w_k y_k \cdot \sum_{k=1}^n w_k x_k^2 - \sum_{k=1}^n w_k x_k y_k \cdot \sum_{k=1}^n w_k x_k \right), \quad (4.16)$$

avec $\Delta = \sum_{k=1}^n w_k x_k^2 \cdot \sum_{k=1}^n w_k - (\sum_{k=1}^n w_k x_k)^2$.

Les incertitudes sur A et B ainsi que la covariance peuvent être déterminées avec les formules du GUM :

$$u^2(A) = \sum_{k=1}^n \left(\frac{\partial A}{\partial y_k} \right)^2 u^2(y_k) = \dots = \frac{1}{\Delta} \sum_{k=1}^n w_k \quad (4.17)$$

$$u^2(B) = \sum_{k=1}^n \left(\frac{\partial B}{\partial y_k} \right)^2 u^2(y_k) = \dots = \frac{1}{\Delta} \sum_{k=1}^n w_k x_k^2 \quad (4.18)$$

$$u(A, B) = \sum_{k=1}^n \frac{\partial A}{\partial y_k} \frac{\partial B}{\partial y_k} u^2(y_k) = \dots = -\frac{1}{\Delta} \sum_{k=1}^n w_k x_k. \quad (4.19)$$

Ces formules paraissent assez complexes mais sont très faciles à programmer. Il faudra bien prendre en compte la corrélation entre A et B par la suite.

4.3.2 Avec une incertitude sur les x_i et y_i

Si l'incertitude est sur les x_i et pas les y_i , ce n'est pas plus compliqué. En effet, il suffit de chercher la droite $x = Ay + B$ qui colle le mieux aux données.

Si l'on cherche à ajuster par une droite $y = Ax + B$ un ensemble de n données (x_i, y_i) qui présentent une incertitude sur les x_i et y_i , alors la plupart des livres proposent d'utiliser la même méthode que précédemment en pondérant par

$$w_k = \frac{1}{u^2(y_k) + A^2 u^2(x_k)}. \quad (4.20)$$

Comme A est une inconnue du problème, il faut alors procéder par itérations : en partant d'une valeur raisonnable de A que l'on injecte dans l'équation précédente, on effectue une régression linéaire par la méthode des moindres carrés qui donne une nouvelle valeur de A qui permet de recommencer la boucle jusqu'à convergence.

Si $y = f(x)$, alors on remplace A par $f'(x_k)$ dans l'équation précédente.

4.3.3 Autres fonctions

On peut faire exactement la même démarche avec un polynôme d'ordre plus élevé. Pour un polynôme d'ordre deux, cela fait 3 coefficients à déterminer. Ce qui est important, c'est la linéarité de la fonction vis à vis des coefficients à déterminer. Ainsi, on peut résoudre facilement les équations issues de la méthode des moindres carrés avec

$$y = A \cos x + B \sin x. \quad (4.21)$$

Pour une fonction exponentielle, $y = Be^{Ax}$, très courante en sciences, on passe par le logarithme pour linéariser le problème :

$$\ln(y) = Ax + \ln B. \quad (4.22)$$

Si la fonction à tester n'est pas linéaire vis à vis des coefficients, on peut commencer la même démarche, mais le système d'équations obtenu ne pourra pas être résolu analytiquement. Il faut alors faire une résolution numérique parfois basée sur un développement limité. On procède alors de manière itérative. Il existe plusieurs algorithmes dans la littérature scientifique dont la description dépasse l'ambition de ce cours. L'article de Donald W. Marquardt [18] sur le sujet est particulièrement cité. On le trouve en ligne gratuitement.

4.3.4 Cas d'une gaussienne

Il est très fréquent de vouloir trouver la gaussienne qui s'ajuste le mieux à un pic. Il existe de nombreux logiciels sur le marché qui procèdent par itération, comme expliqué précédemment. Cependant, il est possible linéariser le problème, comme l'a montré Zimmermann [19] et d'appliquer ensuite une méthode de régression linéaire classique.

On cherche la fonction

$$y(x) = y_0 \exp \left[-\frac{(x - x_0)^2}{2\sigma^2} \right] \quad (4.23)$$

qui colle le mieux à un pic donné caractérisé par les valeurs $Y_i(X_i)$. Pour cela, il faut déterminer les trois paramètres y_0 , x_0 et σ^2 . L'astuce de Zimmermann a été de remarquer que

$$\frac{y(x-1)}{y(x+1)} = \exp\left[\frac{2(x-x_0)}{\sigma^2}\right]. \quad (4.24)$$

Ainsi,

$$A_i = \ln\left[\frac{Y_{i-1}}{Y_{i+1}}\right] = \frac{2}{\sigma^2}(X_i - x_0). \quad (4.25)$$

On peut ensuite effectuer une régression linéaire entre les A_i et les X_i pour déterminer x_0 et σ^2 . Le poids de chaque A_i , w_i , peut être facilement évalué en utilisant la formule de propagation des incertitudes du GUM,

$$w_i = \frac{1}{u^2(A_i)} = \left[\frac{u^2(Y_{i-1})}{Y_{i-1}^2} + \frac{u^2(Y_{i+1})}{Y_{i+1}^2}\right]^{-1}. \quad (4.26)$$

La méthode est testée avec succès par T. Mukoyama [20] qui précise que, pour éviter les distorsions, il vaut mieux éviter les points éloignés du centroïde du pic. Dans son article, Mukoyama suppose que $u^2(Y_i) = Y_i$.

Une fois x_0 et σ^2 déterminés, y_0 est donné par la moyenne pondérée des y_o associés à chaque point

$$\ln(y_0) = \frac{1}{\sum_i w'_i} \sum_i w'_i \left(\ln(Y_i) + \frac{(X_i - x_0)^2}{2\sigma^2} \right), \quad (4.27)$$

avec

$$w'_i = \left[\frac{u^2(Y_i)}{Y_i^2} + \frac{(X_i - x_0)^2}{\sigma^4} \left(u^2(x_0) + \frac{(X_i - x_0)^2}{\sigma^2} u^2(\sigma) + \frac{(X_i - x_0)}{\sigma} u(x_0, \sigma) \right) \right]^{-1}. \quad (4.28)$$

Attention, il manque le terme de covariance dans l'article de Mukoyama.

Cette méthode ne peut pas être appliquée si plusieurs pics se chevauchent.

4.4 Le test du χ^2

4.4.1 Approche intuitive

Ce n'est pas parce que l'on a réussi à choisir la droite qui passe le mieux par un nuage de points, que l'on peut affirmer que ce nuage peut être représenté par une droite. Ou, si l'on hésite entre deux fonctions, comment déterminer celle qui s'ajuste le mieux aux données ?

Pour répondre à cette question, il faut retourner au χ_{min}^2 qui mesure l'écart quadratique à la courbe. Plus il est petit, meilleur sera l'ajustement (le fit). A partir de quelle valeur, considère-t-on l'ajustement acceptable ? Intuitivement, on voit que cela dépend du nombre de termes dans la somme, et donc du nombre de points. On définit le nombre de degrés de liberté, ν , comme étant le nombre de points moins le nombre de paramètres qui ont été fixés. Dans le cas de la régression linéaire vue au paragraphe 4.3.1, $\nu = n - 2$.

Si pour chaque terme de la somme, l'écart entre la courbe et le point est inférieur à un écart-type, alors le χ_{min}^2 obtenu est l'ordre de ν , ou inférieur et l'on peut considérer que l'ajustement est bon. S'il est largement supérieur, ce ne sera pas le cas. On définit parfois le **chi deux réduit** [reduced chi square] par $\tilde{\chi}_{min}^2 = \chi_{min}^2 / \nu$. On compare alors $\tilde{\chi}_{min}^2$ à 1.

4.4.2 Approche quantitative

Un peu de mathématiques pour commencer. Si y est un nombre aléatoire gaussien de moyenne et variance respectives μ , et σ^2 , alors $y - \lambda$ est un nombre aléatoire gaussien de moyenne $\mu - \lambda$ et de variance σ^2 si λ une constante. De même, λy est aussi un nombre aléatoire gaussien de valeur moyenne $\lambda\mu$ et de variance $\lambda^2\sigma^2$. Ainsi, lorsque l'on a atteint le minimum du χ^2 , noté χ_{min}^2 , chaque terme $\frac{y_k - f(x_k)}{\sigma_k}$ est un nombre gaussien de valeur moyenne nulle et de variance égale à 1. Et cela, quelle que soient les données et la fonction étudiée.

En revanche, si y est un nombre aléatoire gaussien, ce n'est plus le cas pour y^2 . La somme de ν carrés de nombres aléatoires gaussiens indépendants de valeur moyenne nulle et de variance égale à 1 satisfait à la **loi de distribution du χ^2** [chi square(d) distribution]. Ainsi,

$$p(\chi_{min}^2, \nu) = \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} [\chi_{min}^2]^{\frac{\nu}{2}-1} \exp\left(-\frac{\chi_{min}^2}{2}\right). \quad (4.29)$$

Cette fonction est représentée figure 4.2. Sa valeur moyenne est ν et sa variance 2ν .

Pour le chi deux réduit, $\tilde{\chi}_{min}^2$, la valeur moyenne est donc égale à 1 et la variance à $2/\nu$.

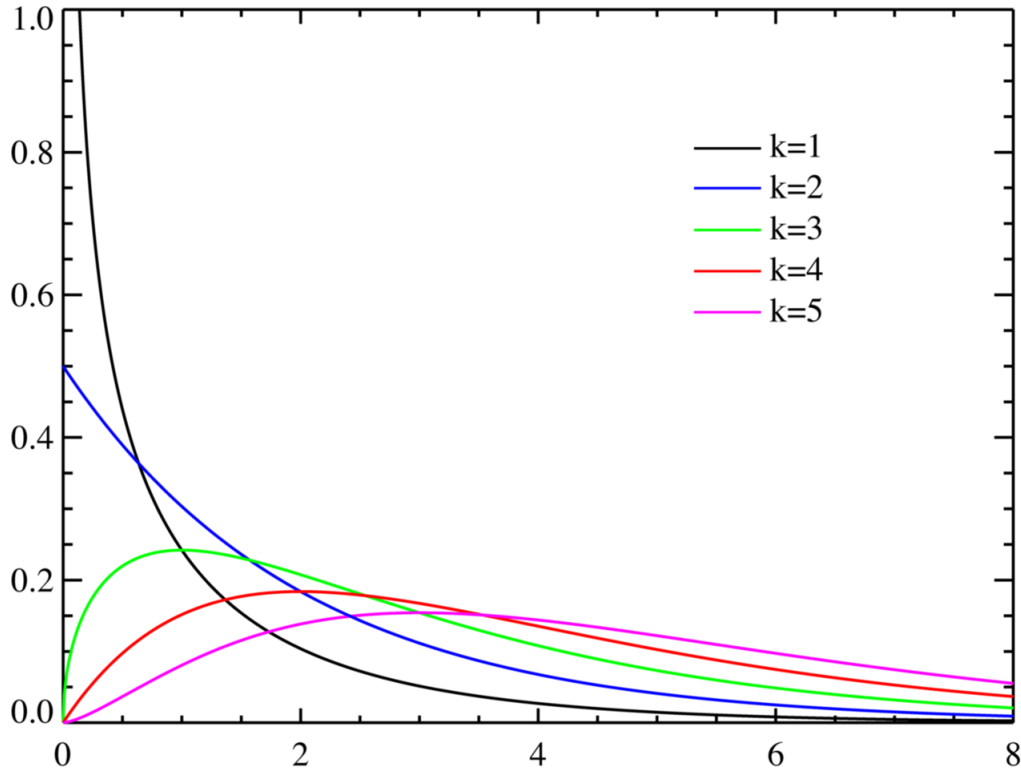


FIG. 4.2 – Fonction de la distribution de probabilité (pdf) du χ^2 pour différents nombres de degrés de liberté k . (Figure tirée de wikipedia.)

Cette pdf permet de calculer la probabilité que le χ^2 soit supérieur au χ_{min}^2 obtenu,

$$Q = \frac{1}{2^{\nu/2} \Gamma(\frac{\nu}{2})} \int_{\chi_{min}^2}^{\infty} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} dx = \frac{1}{\Gamma(\frac{\nu}{2})} \int_{\chi_{min}^2/2}^{\infty} t^{\frac{\nu}{2}-1} e^{-t} dt = \frac{\Gamma(\nu/2, \chi_{min}^2/2)}{\Gamma(\nu/2)}, \quad (4.30)$$

où $\Gamma(a, x) = \int_x^{\infty} t^{a-1} e^{-t} dt$ est la fonction gamma incomplète [incomplete Gamma function]. La fonction Q est tabulée.

Ainsi, si le χ_{min}^2 obtenu est inférieur à $\nu - \sqrt{2\nu}$, Q sera proche de 1 et le fit est très bon. En revanche, si le χ_{min}^2 obtenu est supérieur à $\nu + \sqrt{2\nu}$, Q sera proche de 0 et le fit est mauvais.

Cette loi du χ^2 n'est rigoureusement valable que si les distributions des valeurs y_k sont bien des gaussiennes. Si ce n'est pas le cas, l'information sera plus qualitative.

Il existe de nombreux autres tests du χ^2 pour vérifier des hypothèses statistiques (la distribution est bien gaussienne, deux valeurs sont bien indépendantes, etc). Cela dépasse l'objet de ce cours.

Chapitre 5

Seuil de décision, limite de détection et intervalle de confiance

5.1 Introduction

Une application directe des calculs d'incertitude présentés aux chapitres précédents est la définition du **seuil de décision** [decision threshold], de la **limite de détection** [detection limit] et des **extrémités de l'intervalle de confiance** [limits of the confidence interval] pour lesquels une norme internationale a été publiée en 2009 (mai 2010 pour la version française) : *Détermination des limites caractéristiques (seuil de décision, limite de détection et extrémités de l'intervalle de confiance) pour mesurages de rayonnements ionisants - Principes fondamentaux et applications*, NF ISO 11929.

Cette norme est difficile à lire et à comprendre, mais l'article suivant est très utile à la compréhension : A. Vivier et al. *Radioprotection* 45 (2010) 345. Je vais donc commencer par des situations où les équations sont très simples afin de comprendre les concepts.

Quel est le problème? Un signal expérimental n'est pas toujours pur. Il peut être "pollué" par un bruit de fond. Sur la figure 5.1, où le pic est quasiment inexistant, y a-t-il un contaminant ou est-ce une simple fluctuation du bruit de fond? Dans le cas où l'on est convaincu qu'il y a bien des événements significatifs, quelle est alors l'incertitude?

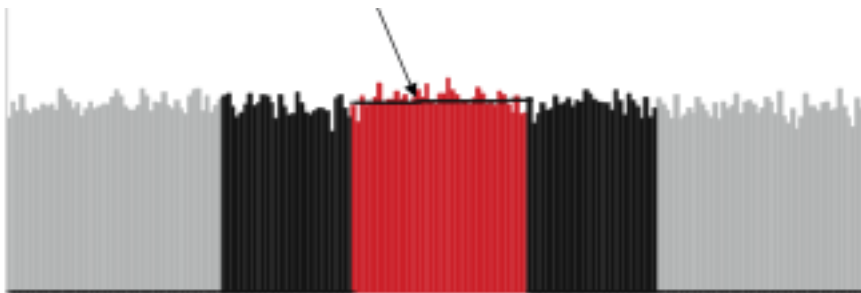


FIG. 5.1 – Recherche d'un pic dans un spectre. (figure tirée de l'article de Vivier et al. [7])

Si un pic voisin vient interférer avec le pic à analyser, cela complique encore plus l'analyse. Nous allons supposer ici, que le pic recherché est seul et vient s'ajouter au bruit de fond.

La nouvelle norme de 2009 se distingue des anciennes normes car elle utilise les méthodes dites bayésiennes qui seront vues plus en détail au chapitre suivant. Ces méthodes sont pertinentes quand il y a peu de données et qu'il est utile de prendre en compte d'autres informations connues de l'expérimentateur pour corriger les résultats. Par exemple, si l'on trouve (2 ± 1) Bq, il y a une probabilité non négligeable d'avoir des valeurs négatives qui n'ont pas de sens physique. Quand le signal expérimental est très faible, il faut alors le corriger par des certitudes (l'activité est positive). La norme répond aussi à ce problème.

J'ai pris le parti ici de commencer par l'application pratique avant d'expliquer les concepts sous jacents dans leur détail.

Pour répondre à toutes ces questions, il faut connaître les lois de distribution statistique pour le bruit de fond et pour le signal recherché. En faisant l'hypothèse que ces lois satisfont à la statistique de Poisson, le calcul est assez simple. Si on ne la connaît pas, les expressions mathématiques sont plus compliquées et la résolution de fait de façon approximative.

5.2 Avec une distribution de Poisson

5.2.1 Distribution de Poisson

Quand on compte un nombre d'évènements, on ne doit pas faire d'erreur et l'on peut généralement supposer qu'il n'y a pas d'incertitude de mesure sur le comptage. Sauf bien-sûr si le détecteur sature par exemple. Ici, on supposera qu'il n'y a pas d'incertitude de comptage. En revanche, le processus que l'on étudie peut avoir un caractère aléatoire. Si l'on fait plusieurs comptages de suite, on aura des valeurs différentes. C'est le cas en particulier pour la radioactivité.

La distribution de Poisson donne la probabilité de compter la valeur n pour un nombre d'occurrences moyen égal à λ ,

$$P(n, \lambda) = \frac{\lambda^n}{n!} e^{-\lambda}. \quad (5.1)$$

Ici, n est un entier, mais pas forcément λ . $n!$ est la factorielle de n . La fonction $P(n, \lambda)$ est normalisée à 1. Son allure est tracée figure 5.2.

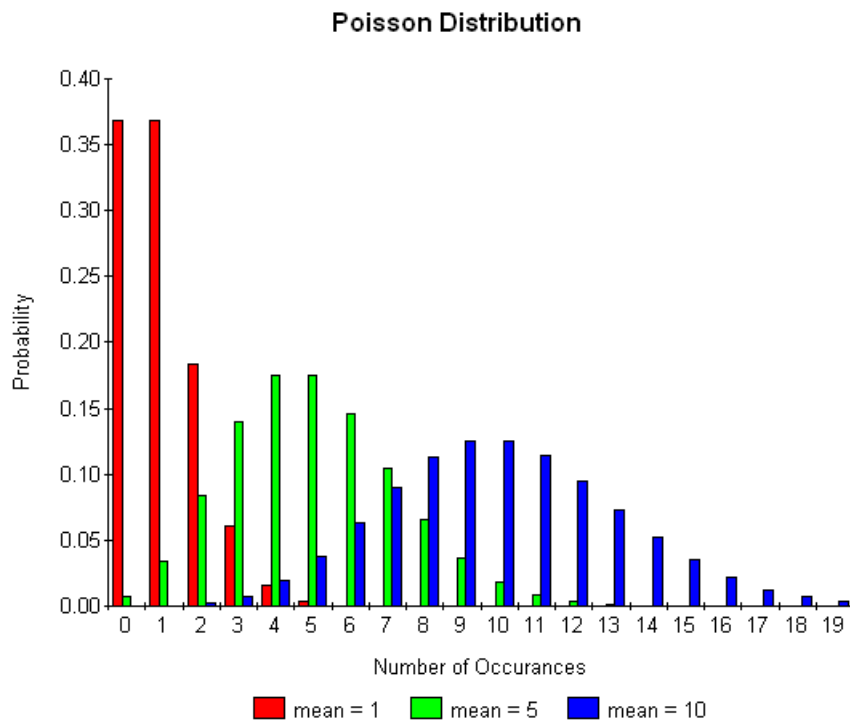


FIG. 5.2 – Distribution de Poisson (figure tirée de `systemath.com`).

La valeur moyenne de la distribution de Poisson peut être aisément calculée et est égale à λ . La variance de la distribution est aussi égale à λ et l'écart-type vaut donc $\sqrt{\lambda}$.

La démonstration se fait avec le développement en série de la fonction exponentielle. Pour la moyenne, c'est rapide,

$$\bar{n} = \sum_{n=1}^{\infty} nP(n, \lambda) = e^{-\lambda} \sum_{n=1}^{\infty} \frac{n\lambda^n}{n!} = e^{-\lambda} \lambda \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} = e^{-\lambda} \lambda e^{\lambda} = \lambda. \quad (5.2)$$

Pour la variance, c'est un peu plus long, mais pas plus compliqué :

$$\overline{n^2} = \sum_{n=1}^{\infty} n^2 P(n, \lambda) = e^{-\lambda} \lambda \sum_{n=1}^{\infty} n \frac{\lambda^{n-1}}{(n-1)!} \quad (5.3)$$

$$= e^{-\lambda} \lambda \left(\sum_{n=1}^{\infty} (n-1) \frac{\lambda^{n-1}}{(n-1)!} + \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} \right) \quad (5.4)$$

$$= e^{-\lambda} \lambda (\lambda e^{\lambda} + e^{\lambda}) = \lambda^2 + \lambda = (\bar{n})^2 + \lambda. \quad (5.5)$$

Quand $\lambda \rightarrow \infty$, la distribution de Poisson tend vers une distribution normale de valeur moyenne λ et de variance λ . En fait, dès que $\lambda > 9$, la distribution de Poisson est très proche d'une distribution gaussienne. On peut donc utiliser ce que nous connaissons sur les gaussiennes en terme de facteur d'élargissement pour calculer des probabilités.

5.2.2 Seuil de décision

Définition

Imaginons que l'on fasse un comptage par spectrométrie et que le résultat ressemble à celui de la figure 5.1. L'abscisse correspond au numéro de canal ou à l'énergie et l'ordonnée au nombre de coups. Supposons que le bruit de fond ait une valeur $n_{bdf} = 46$ et que le résultat brut soit $n_{brut} = 54$. Cela fait un résultat net de $n_{net} = n_{brut} - n_{bdf} = 8$. On ne peut pas affirmer que ces 8 coups supplémentaires sont dus au signal (radioactivité) que l'on cherche. En revanche, si n_{net} est grand, il est beaucoup moins probable que ces événements supplémentaires soient dues à une fluctuation du bruit de fond. Où est la limite entre ces deux exemples extrêmes ? A partir de quand une mesure devient significative ? C'est ce que l'on appelle le seuil de décision défini ainsi dans la norme :

Seuil de décision : valeur de l'estimateur du mesurande telle que, quand le résultat d'une mesure réelle utilisant une procédure de mesure donnée d'un mesurande quantifiant le phénomène physique lui est supérieur, on décide que le phénomène physique est présent.

Ce sera sûrement plus clair quand nous l'aurons calculé.

Calcul

Pour trouver le seuil de décision, on va supposer que la valeur vraie de n_{net} , $\mu_{net} = 0$. Comme $n_{net} = n_{brut} - n_{bdf}$,

$$u^2(\mu_{net} = 0) = u^2(n_{brut}) + u^2(n_{bdf}) = n_{brut} + n_{bdf} = 2n_{bdf}. \quad (5.6)$$

Le problème avec ce résultat, c'est que si $n_{bdf} = 0$ les variances sont nulles. Cela n'est pas réaliste car, avec une durée de mesurage finie, on ne peut jamais être sûr que $n_{bdf} = 0$. Comme suggéré dans la norme, ce défaut, par exemple lors de comptages à faible niveau, peut être évité en remplaçant tous les résultats de comptage n par $n + 1$. Ce n'est pas du bricolage, le terme $+1$ apparaît dans une détermination rigoureuse par une méthode bayésienne comme nous le démontrerons dans le chapitre suivant. Ainsi,

$$u^2(\mu_{net} = 0) = 2(n_{bdf} + 1). \quad (5.7)$$

La distribution de valeurs possibles pour n_{net} a donc une valeur moyenne nulle et un écart-type égal à $\sqrt{2(n_{bdf} + 1)}$, voir figure 5.3.

Pour une valeur de n_{net} plus petite que $u(\mu_{net} = 0)$, la probabilité qu'il s'agisse d'une fluctuation du bruit de fond est grande. Inversement, si $n_{net} \gg u(\mu_{net} = 0)$, cette même probabilité sera très faible. Pour définir la frontière entre ces deux cas, on se fixe une valeur de risque de se tromper α en affirmant que le signal est présent. Cette valeur est généralement choisie assez faible.

Le seuil de décision (SD) est donc défini à partir d'un **intervalle de risque** α par

$$P(n_{net} > SD; \bar{n}_{net} = 0) = \alpha. \quad (5.8)$$

Cette équation peut être lue comme étant que la probabilité que n_{net} soit supérieur au seuil de détection est égale à α . Ou, inversement, que n_{net} soit inférieur à SD est égale à $1 - \alpha$. α est généralement choisi petit, entre 2,5% et 5%.

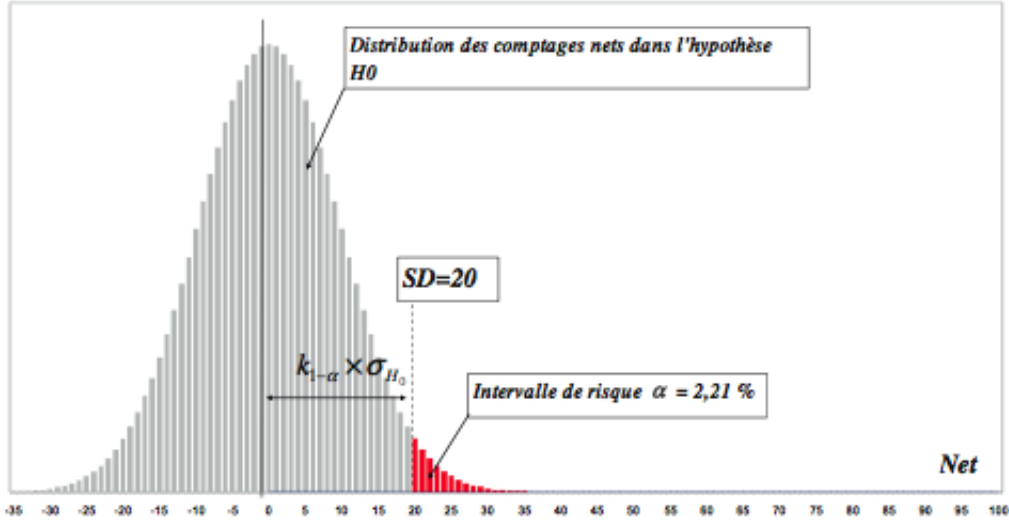


FIG. 5.3 – Exemple de distribution de n_{net} et de seuil de décision (figure tirée de Vivier et al. [7]).

Finalement, avec une statistique de Poisson, le seuil de décision est simplement égal à

$$SD = k_{1-\alpha} \sqrt{2(n_{BDF} + 1)}, \quad (5.9)$$

où $k_{1-\alpha}$ est le facteur d'élargissement correspondant à la probabilité $1 - \alpha$.

Attention, ce facteur d'élargissement n'est pas le même que celui utilisé pour exprimer l'incertitude ! Il suffit de comparer les figures 2.2 et 5.3 pour comprendre. En langage courant, deux ailes de la gaussienne sont rognées dans la figure 2.2 et une aile dans la figure 5.3. Dans le chapitre précédent, un facteur d'élargissement de 2 donne 95% de chance d'être dans l'intervalle intérieur. Ce qui donne 5% de chance d'être dans les ailes et donc 2,5% de chance d'être dans l'aile droite.

Pour une distribution gaussienne et un nombre infini de degrés de liberté, le lien entre α et $k_{1-\alpha}$ est donné par

$$\alpha = \frac{1}{2} \operatorname{erfc} \left[\frac{k_{1-\alpha}}{\sqrt{2}} \right] \quad \text{avec} \quad \operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-y^2} dy. \quad (5.10)$$

La fonction erreur complémentaire étant difficile à inverser, les normes proposent des tables.

L'hypothèse faite ici est celle d'une distribution gaussienne. Pour n_{bdf} suffisamment élevé, la distribution de Poisson tend vers une gaussienne. Voir figure 5.2. En revanche, cette hypothèse n'est plus correcte pour n_{bdf} petit. La frontière se trouve vers $n_{bdf} = 10$. En effet, pour $n_{bdf} > 10$, l'écart entre l'approche gaussienne et une approche bayésienne plus rigoureuse est inférieur à 5%, ce qui est plus petit que l'incertitude sur le seuil de décision qui est de 14% pour $n_{bdf} = 10$. Ce point est discuté par exemple dans la référence [21].

Finalement, on a $\alpha\%$ chances d'obtenir une valeur supérieure à SD si l'on mesure un échantillon réellement non radioactif. Ou, dit autrement, si on mesure N échantillons non radioactifs, $\alpha\%$ d'entre eux auront une valeur nette supérieure à SD .

Mise en œuvre

En éliminant le cas limite où $n_{net} = SD$, il reste deux cas possibles : $n_{net} > SD$ ou $n_{net} < SD$.

Premier cas : Si $n_{net} > SD$, alors la probabilité que l'échantillon soit non-radioactif est inférieure à $\alpha\%$. On la rejette donc et l'on considère la mesure comme significative d'une radioactivité avérée. L'activité de l'échantillon (nombre de désintégrations par seconde) est alors déduite de n_{net} ,

$$A = \frac{n_{net}}{\epsilon t}, \quad (5.11)$$

où t est le temps de comptage et ϵ , l'efficacité du détecteur.

Reprenons l'exemple du début avec $n_{bdf} = 46$. Cela correspond à un $SD = 20$ avec $\alpha = 2,5\%$ ($k_{1-\alpha} = 1,96 \simeq 2$). Il faut donc que $n_{net} > 20$ et $n_{brut} = 46 + 20 = 66$ pour que le comptage soit considéré comme significatif avec une probabilité supérieure à 97,5%. Plus n_{net} sera élevé moins on aura de chances de se tromper en déclarant l'échantillon radioactif.

Deuxième cas : Si $n_{net} < SD$, il est possible que l'échantillon ne soit pas radioactif. Mais on ne peut pas rejeter l'hypothèse que l'échantillon est radioactif. Une très faible activité réellement présente sera masquée par les fluctuations du bruit de fond et peut conduire à $n_{net} < SD$.

Dans ce cas, tout ce que l'on peut affirmer, c'est que la radioactivité présente dans l'échantillon est inférieure à une valeur appelée limite de détection que l'on détermine par la suite.

Conclusion : La définition du seuil de détection dans la norme NF ISO 11929 est assortie de deux notes reprises ci-dessous.

NOTE 1 : Le seuil de décision est défini de manière que, dans le cas où le résultat du mesurage, y , dépasse le seuil de décision, y^ , la probabilité que la valeur vraie du mesurande soit nulle est inférieure ou égale à la probabilité choisie α .*

NOTE 2 : Si le résultat, y , est inférieur au seuil de décision, y^ , le résultat ne peut pas être attribué à l'effet physique ; néanmoins, il ne peut être conclu qu'il est absent.*

5.2.3 Limite de détection

Définition

La norme internationale définit la la limite de détection ainsi :

Limite de détection : *Plus petite valeur vraie du mesurande qui garantit une probabilité spécifiée qu'il soit détectable par la méthode de mesure.*

Cette définition est un peu plus limpide que la précédente.

Calcul

Dans la section précédente, on est parti de l'hypothèse qu'il n'y avait pas de signal ($\mu_{net} = 0$) et l'on a étudié la statistique de cette hypothèse pour calculer la probabilité qu'elle soit vraie ou fausse. Maintenant, on va supposer qu'il y a un signal ($n_{net} > 0$) et étudier sa statistique pour déterminer jusqu'à quelle valeur il est indétectable.

La démarche va être similaire à celle adoptée pour le seuil de détection. On considère un échantillon dont on connaît parfaitement la valeur moyenne vraie du comptage net qu'il génère, μ_{net} . On construit ensuite la distribution des valeurs n_{net} associée à cette valeur moyenne et à sa variance $u^2(\mu_{net})$. Graphiquement, cela correspond à la figure 5.4. Le pic de gauche (H0) correspond à l'hypothèse $\mu_{net} = 0$ et est identique à la figure 5.3. Celui de droite (H1) correspond à l'hypothèse étudiée maintenant où $\mu_{net} > 0$.

Si la valeur vraie de du nombre de coups $\mu_{net} = SD$, alors on a 50% de chances de déclarer que l'on n'a pas vu de radioactivité dans cet échantillon et 50% de chances de déclarer l'échantillon est radioactif. Or l'échantillon est bel et bien radioactif! Si l'on augmente la valeur de μ_{net} , ces pourcentages vont changer : le premier va diminuer et le second va augmenter.

On cherche ensuite la valeur de μ_{net} , appelée LD , qui conduit $\beta\%$ des cas à avoir $n_{net} < SD$. Ce qui signifie que l'on a $\beta\%$ de chances de déclarer cet échantillon non radioactif, alors qu'il l'est réellement. Ce risque d'erreur est appelé *erreur de seconde espèce ou risque β* .

On déduit aisément du graphique de la figure 5.4 que

$$LD = SD + k_{1-\beta}u(\mu_{net} = LD). \quad (5.12)$$

La difficulté de cette relation, c'est que $u(\mu_{net} = LD)$ dépend de LD ! En effet, $\mu_{net} = n_{brut} - n_{bdf}$ et donc pour une distribution de Poisson,

$$u^2(\mu_{net}) = u^2(n_{brut}) + u^2(n_{bdf}) = (n_{brut} + 1) + (n_{bdf} + 1) = 2(n_{bdf} + 1) + \mu_{net}. \quad (5.13)$$

Finalement, dans le cas présent, LD est défini par

$$LD = SD + k_{1-\beta}\sqrt{2(n_{bdf} + 1) + LD}. \quad (5.14)$$

La résolution se fait généralement de manière itérative ou approximative. Plusieurs exemples sont traités dans la norme internationale.

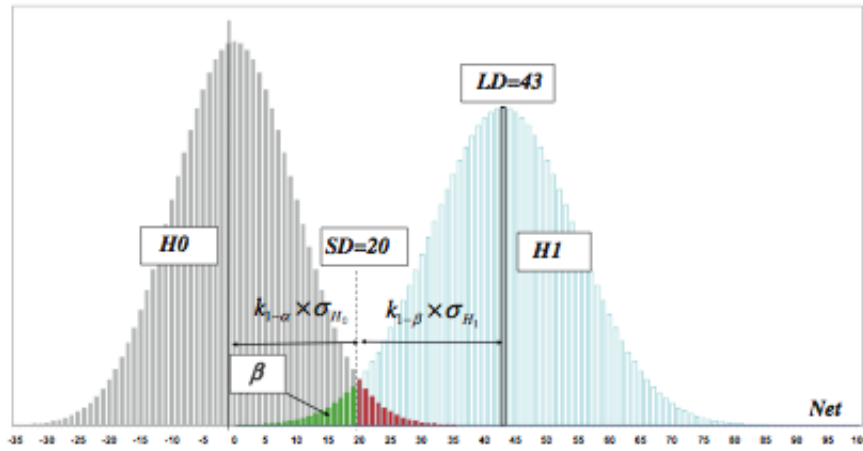


FIG. 5.4 – Exemple de distribution de n_{net} et de limite de détection (figure tirée de Vivier et al. [7]).

En fixant $\alpha = \beta$, on a $k_{1-\alpha} = k_{1-\beta}$ et cela simplifie les expressions. C'est ce que nous ferons à partir de maintenant.

On peut remarquer qu'en annulant LD dans la racine de l'équation (5.14), on obtient $LD = 2SD$. Ainsi, on a toujours $LD > 2SD$.

Résolution

A partir de la définition de SD et en supposant toujours que $k_{1-\alpha} = k_{1-\beta} = k$ pour alléger les notations, on peut réécrire l'équation (5.14) ainsi,

$$LD = SD + k\sqrt{\frac{SD^2}{k^2} + LD}. \quad (5.15)$$

Parce que le modèle est simple ici, on peut le résoudre exactement pour trouver

$$LD = 2SD + k^2. \quad (5.16)$$

Cette équation est souvent utilisée pour la radioactivité. Comme k est de l'ordre de 2, voire moins, si SD est suffisamment grand, on a finalement le résultat bien connu et souvent appliqué de $LD \simeq 2SD$.

Comme, dans les cas plus complexes, on ne peut pas résoudre analytiquement l'équation (5.15), voici deux approches possibles.

Résolution approximative : En remarquant que le premier terme dans la racine est grand devant le second, on peut faire un développement limité et aboutir à

$$LD \simeq 2SD + k^2. \quad (5.17)$$

C'est le résultat exact ! Un développement limité ne mène pas toujours à la solution exacte !

Reprenons l'exemple numérique du début. Nous avons $n_{BDF} = 46$ et $SD \simeq 20$ avec $k = 2$. Ainsi, $LD \simeq 44$.

Résolution itérative : On part de $LD_0 = 2SD = 40$. On injecte cette valeur dans le membre de droite de l'équation (5.14) ou (5.15) et l'on obtient $LD_1 = 43,66$. On recommence le processus avec LD_1 et l'on obtient $LD_2 = 43,97$. $LD_3 = 44,00$. Cela converge assez vite.

Mise en œuvre

Rappelons que la limite de détection n'est utile que dans le cas où $n_{net} < SD$. On a $\beta\%$ chances d'obtenir une valeur inférieure à SD si $n_{net} = LD$. Ainsi, si $n_{net} < SD$ la probabilité que la valeur vraie de n_{net} soit supérieure à LD est inférieure à $\beta\%$.

Quand le résultat d'un mesurage conduit à $n_{net} < SD$, alors on affiche qu'il est inférieur à la limite de détection. De fait, l'intervalle $[0, LD]$ correspond à l'intervalle de confiance au niveau $1 - \beta\%$ de la valeur vraie. La limite de détection n'est donc pas une valeur probable, mais la limite supérieure de l'intervalle de confiance.

La limite de détection est souvent utilisée pour quantifier la performance d'une mesure ou d'un laboratoire.

5.2.4 Calcul de l'activité

Les résultats d'un mesurage ne sont pas exprimés en nombre de coups, mais dans une grandeur physique. Pour une activité, la relation est

$$A = \frac{n_{net}}{\epsilon \cdot t}, \quad (5.18)$$

où t est le temps de comptage et ϵ , l'efficacité du détecteur. La mesure du temps est généralement suffisamment précise pour que l'on puisse négliger son incertitude. En revanche, ce n'est pas toujours le cas pour l'efficacité du détecteur qui prend en compte la géométrie de la source et sa distance au détecteur. . .

Le seuil de décision ne dépend que du bruit de fond et n'est donc pas affecté par ces incertitudes supplémentaires. En effet, l'activité de départ est nulle, quelle que soit l'efficacité du détecteur.

Ainsi, le seuil de décision en terme d'activité, SD_A , s'écrit

$$SD_A = \frac{SD}{t \cdot \epsilon}. \quad (5.19)$$

En effet, si $n_{net} > SD$, alors $\frac{n_{net}}{\epsilon \cdot t} > \frac{SD}{\epsilon \cdot t}$ et donc $A > SD_A$.

En revanche, une incertitude élevée sur la valeur du rendement affecte la limite de détection exprimée en terme d'activité. A partir de l'équation (5.18) qui est un produit, on a immédiatement que

$$\frac{u^2(A)}{A^2} = \frac{u^2(n_{net})}{n_{net}^2} + \frac{u^2(\epsilon)}{\epsilon^2} \quad (5.20)$$

et l'équation (5.14) est remplacée par,

$$LD_A = SD_{[Bq]} + k \times u(LD_A). \quad (5.21)$$

Finalement, après quelques lignes de calcul, on obtient,

$$LD = \frac{k^2 + 2SD}{1 - u_{\%}^2(\epsilon) \times k^2} \quad \text{et} \quad LD_A = \frac{LD}{\epsilon \cdot t}, \quad (5.22)$$

où $u_{\%}(\epsilon) = \frac{u(\epsilon)}{\epsilon}$ est l'incertitude relative sur le rendement.

Avec $u_{\%}(\epsilon) = 30\%$ et $\alpha = \beta = 2,5\%$, la prise en compte de l'incertitude sur le rendement multiplie la limite de détection par 1,56. Dans ces conditions, $LD \simeq 3SD$.

En conclusion, le seuil de décision est défini par rapport à la valeur nulle, que ce soit pour l'intégrale nette ou l'activité. L'incertitude sur l'efficacité du détecteur n'intervient pas. En revanche, la valeur de la limite de détection doit prendre en compte cette incertitude.

5.2.5 Extrémités de l'intervalle de confiance et meilleure estimation

Si le résultat du mesurage est supérieur au seuil de décision, alors, l'effet physique est reconnu comme présent. Si ce résultat est proche du seuil et que l'on adopte la démarche habituelle pour déterminer l'intervalle de confiance, il se peut que l'on obtienne des valeurs négatives. Voir la figure 5.5. Or, dans de nombreux cas, le mesurande ne peut pas être négatif. C'est le cas pour la radioactivité.

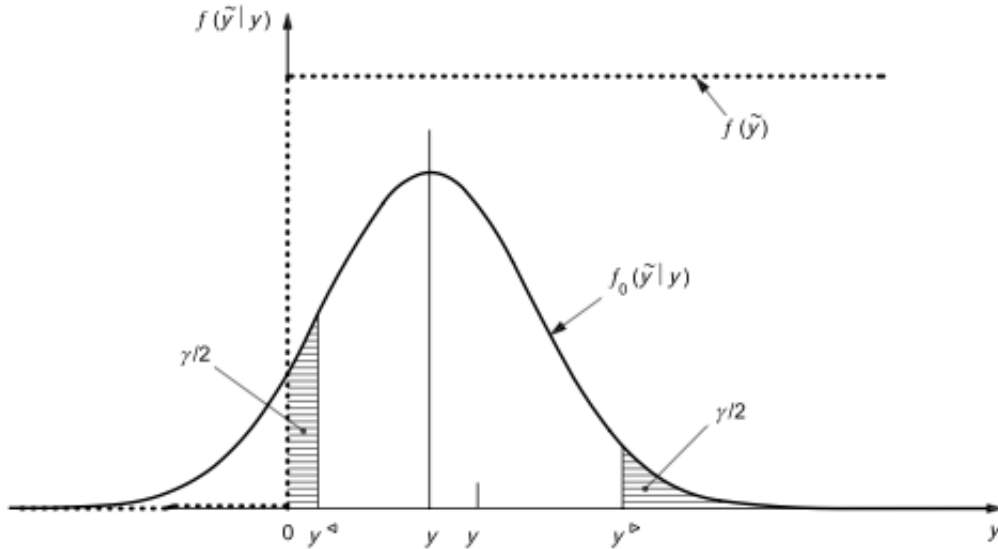


FIG. 5.5 – Définition des extrémités de l'intervalle de confiance et de la meilleure estimation (figure tirée de la norme ISO 11929 :2010(F)).

Comme on peut aussi bien utiliser le résultat de mesure primaire du mesurande, n_{net} ou A , l'activité, on notera y les valeurs possibles du mesurande et $u(y_0)$ l'incertitude type associée ($u(n_{net})$ ou $u(A)$). Les limites inférieure y^{\inf} et supérieure y^{\sup} de l'intervalle de confiance sont définies à partir de la probabilité que la valeur vraie de y soit comprise dans l'intervalle $[y^{\inf}, y^{\sup}]$ est égale à $1 - \gamma$:

$$\frac{\gamma}{2} = \int_0^{y^{\inf}} f(y) dy = \int_{y^{\sup}}^{+\infty} f(y) dy. \quad (5.23)$$

Veillez noter que les bornes ne sont pas symétriques dans les deux cas, conformément à ce qui est représenté figure 5.5.

A partir d'une distribution normale pour $f(y) = \frac{1}{\sqrt{2\pi}u(y_0)} \exp\left[-\frac{(y-y_0)^2}{2u^2(y)}\right]$, on a les relations suivantes :

$$\frac{\gamma}{2} = \frac{1}{2} \left(\operatorname{erfc} \left[\frac{-y_0}{\sqrt{2}u(y_0)} \right] - \operatorname{erfc} \left[\frac{y^{\inf} - y_0}{\sqrt{2}u(y_0)} \right] \right) \quad (5.24)$$

$$\frac{\gamma}{2} = \frac{1}{2} \operatorname{erfc} \left[\frac{y^{\sup} - y_0}{\sqrt{2}u(y_0)} \right], \quad (5.25)$$

qui ne sont pas facile à inverser. La norme propose donc d'écrire le résultat en terme de facteur d'élargissement qui sont tabulés et mieux connus :

$$y^{\inf} = y_0 - k_p u(y_0) \quad \text{avec} \quad p = \omega(1 - \gamma/2) \quad (5.26)$$

$$y^{\sup} = y_0 + k_q u(y_0) \quad \text{avec} \quad q = 1 - \omega\gamma/2. \quad (5.27)$$

Le paramètre ω est donné par

$$\omega = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_0/u(y_0)} \exp\left[-\frac{x^2}{2}\right] dx. \quad (5.28)$$

ω est aussi tabulé dans la norme. Si $y_0 \gg u(y_0)$ ($y_0 > 4u(y_0)$ suffit) alors $\omega \rightarrow 1$ et $p = q = \gamma/2$. On retrouve un résultat simple. Mais pour de faibles comptages proches du seuil de décision, les expressions ci-dessus doivent être utilisées, même si elles sont plus compliquées.

Comme l'intervalle de confiance n'est plus symétrique, voir figure 5.5, la valeur moyenne n'est plus la valeur la plus probable. La norme donne

$$\bar{y} = y_0 + \frac{u(y_0)}{\omega\sqrt{2\pi}} \cdot \exp\left[-\frac{y_0^2}{2u^2(y_0)}\right], \quad (5.29)$$

avec, pour incertitude associée,

$$u(\bar{y}) = \sqrt{u^2(y_0) - (\bar{y} - y_0)\bar{y}}. \quad (5.30)$$

La démonstration n'est pas immédiate, je ne la ferai donc pas. Là encore, si $y_0 > 4u(y_0)$, on retrouve $\bar{y} \simeq y_0$ et $u(\bar{y}) \simeq u(y_0)$.

5.3 Exemples d'optimisation du seuil de décision

Cette partie est directement inspirée de l'article de Vivier et al [7] mentionné dans l'introduction. En considérant toujours la recherche de contaminants radioactifs dans un échantillon comme exemple, on distingue deux cas. Si l'échantillon ne contribue pas au bruit de fond, la mesure du bruit de fond peut se faire en l'absence de l'échantillon et peut par conséquent être plus longue que pour la recherche d'un contaminant. Nous allons rechercher le rapport optimal des temps de mesure. En revanche, si l'échantillon contribue au bruit de fond, il faut extraire le bruit de fond sur le spectre même. Nous étudierons ce cas aussi.

5.3.1 Influence des durées de comptage

Supposons que le bruit de fond puisse être mesuré sans la présence des échantillons à analyser sur des temps donc différents des temps de comptage. En notant t_{bdf} le temps de comptage du bruit de fond, t_s le temps de comptage des échantillons et $\alpha = t_{bdf}/t_s$, le bruit de fond à retrancher au comptage brut est

$$\bar{n}_{bdf} = n_{bdf} \frac{t_s}{t_{bdf}} = \frac{n_{bdf}}{\alpha}. \quad (5.31)$$

L'incertitude sur le bruit de fond mesuré ainsi est

$$u^2(\bar{n}_{bdf}) = \frac{1}{\alpha^2} u^2(n_{bdf}) = \frac{1}{\alpha^2} (n_{bdf} + 1) = \frac{\bar{n}_{bdf}}{\alpha} + \frac{1}{\alpha^2}. \quad (5.32)$$

On voit là tout l'intérêt de mesurer le bruit de fond sur des temps longs : l'incertitude sera d'autant plus petite que α est grand.

Comme $n_{net} = n_{brut} - \bar{n}_{bdf}$ et que n_{brut} est compté sur un temps correspondant à $\alpha = 1$, on a finalement,

$$u^2(n_{net=0}) = u^2(n_{brut}) + u^2(\bar{n}_{bdf}) = \bar{n}_{bdf} + 1 + \frac{\bar{n}_{bdf}}{\alpha} + \frac{1}{\alpha^2}. \quad (5.33)$$

Un calcul plus rigoureux basé sur une approche bayésienne [21] donne

$$u^2(n_{net=0}) = \left(1 + \frac{1}{\alpha}\right) \left(\bar{n}_{bdf} + \frac{1}{\alpha}\right). \quad (5.34)$$

Quelle que soit la formule adoptée, le seuil de décision, peut être globalement divisé par $\sqrt{2}$ pour α très grand, c'est à dire, pour un temps de comptage du bruit de fond beaucoup plus grand que le temps de comptage des échantillons à analyser.

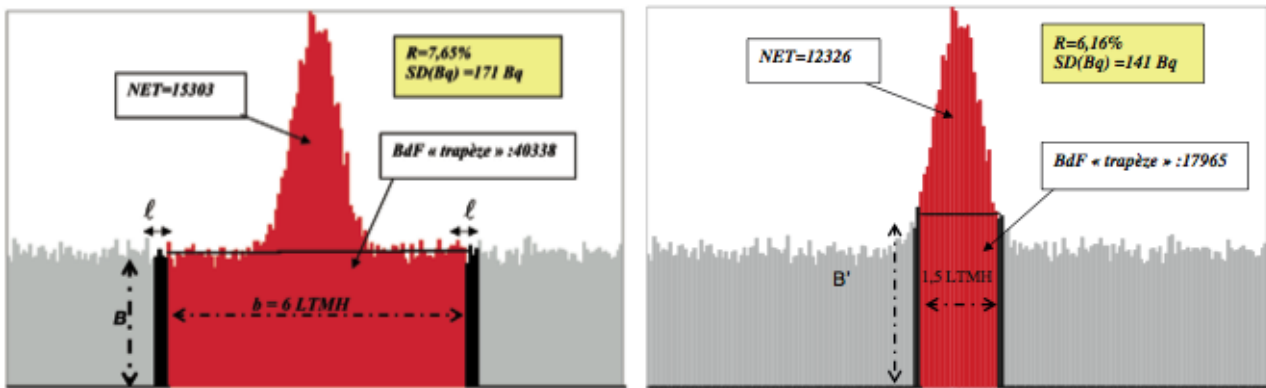


FIG. 5.6 – Exemples extrêmes de sélection de la zone d'intérêt pour extraire le bruit de fond d'un pic (figure tirées de Vivier et al [7]).

5.3.2 Extraction du bruit de fond d'un spectre

Si l'échantillon contribue au bruit de fond, on ne peut pas le mesurer séparément et il faut le retrancher directement du spectre. La figure 5.6 représente deux cas extrêmes. A gauche, la zone d'intérêt [region of interest] est très large. n_{bdf} sera donc grand et le seuil de détection élevé. A droite, la zone d'intérêt est étroite. Le bruit de fond inclut des événements qui appartiennent manifestement au signal recherché.

Un calcul de largeur optimale peut être fait de façon à avoir le seuil de décision le plus faible. On peut montrer que lorsque le pic est gaussien, la zone d'intérêt optimale a une largeur égale à 2,5 fois la largeur à mi-hauteur [FWHM] [7]. Dans la pratique, il est conseillé de prendre 3 fois cette largeur.

5.4 Cas général

Les expressions données pour le seuil de décision et la limite de détection ont été calculées avec l'hypothèse d'une loi de Poisson. Cela ne s'applique pas à tous les problèmes. La norme propose des formules beaucoup plus générales. Ce qui est important, c'est de comprendre la démarche et de pouvoir l'appliquer aux cas spécifiques qui vous intéressent. Je ne vais donc pas reprendre ces formules ici et je vous renvoie à la norme si vous voulez en savoir plus.

5.5 Limite de décision

En plus du seuil de décision et de la limite de détection, il existe aussi la **limite de décision** [decision limit] qui est utilisée pour décider de la **conformité** [compliance] d'un produit. Pour la chimie, il existe un guide en anglais intitulé *Use of uncertainty information in compliance assessment* (2007) commun à Eurachem et CITAC qui peut être téléchargé gratuitement sur : <http://www.measurementuncertainty.org/>. Il a inspiré cette partie.

Quand on veut vérifier si un résultat de mesure respecte ou non un cahier des charges [specification], il est nécessaire de prendre en compte l'incertitude de mesure. La figure 5.7 montre quelques scénarii.

Les cas (i) et (iv) sont assez clairs. En revanche, pour décider de la conformité ou non dans les cas (ii) et (iii), il est nécessaire de définir des **règles de décision** [decision rules] qui définissent la probabilité acceptable de prendre la mauvaise décision. A partir de cette règle et de l'incertitude du mesurage, on peut définir une **plage de conformité** [acceptation zone] et une **plage de non conformité** [rejection zone] .

Ainsi, la plage de non conformité commence à la limite de conformité L plus ou moins une valeur g ([guard band]) qui est choisie de façon à ce que la probabilité de rejet par erreur soit inférieure à α si le résultat est supérieur à $L \pm g$ (le signe dépend de la situation et de la règle).

g est en général un multiple de l'incertitude type u . Pour une distribution normale ou proche, $g = 1,64 u$ donne $\alpha = 5\%$ et $g = 2,33 u$, $\alpha = 1\%$. Une incertitude élevée signifie une plage de conformité plus étroite. Voir

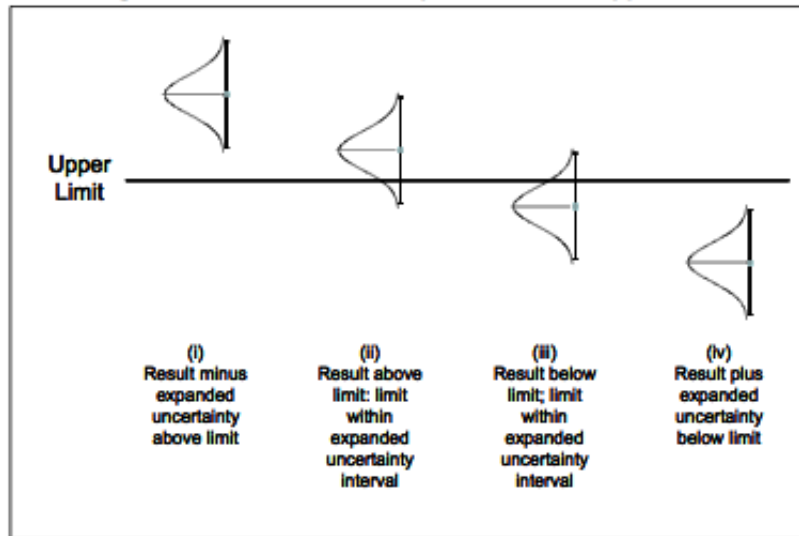


FIG. 5.7 – Vérification du respect d'un cahier des charges avec une limite supérieure (figure tirée du guide *Use of uncertainty information in compliance assessment* (2007) commun à Eurachem et CITAC).

la figure 5.8. Par conséquent, le nombre d'échantillons rejetés sera plus grand. Mais, une valeur plus faible de u a un coût. L'incertitude idéale est souvent choisie de façon à minimiser les coûts.

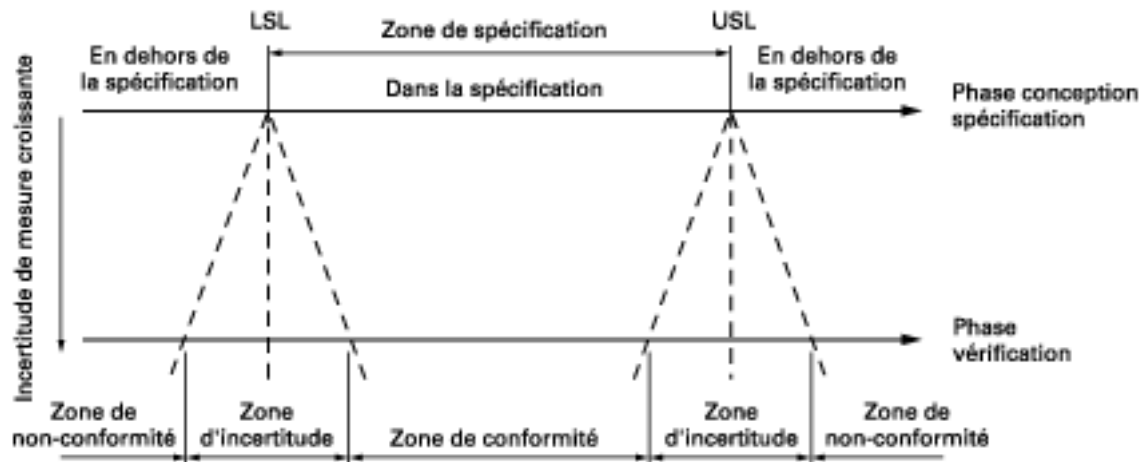


FIG. 5.8 – Illustration d'une stratégie de décision (figure tirée de la norme FD X 07-021)

D'une manière générale, le calcul de g peut être aussi compliqué que pour le seuil de décision ou la limite de détection. Le guide donne quelques exemples que je ne reprendrai pas.

Chapitre 6

Méthodes bayésiennes de calcul de l'incertitude

« *La théorie des probabilités n'est que le bon sens réduit au calcul.* »

Pierre Simon, Marquis de Laplace

Lors de l'introduction de ce cours, nous sommes partis de la définition de l'espérance et de la variance au sens mathématique. Le but de la mesure est d'estimer au mieux ces deux grandeurs. Pour cela on multiplie le nombre de mesurages et l'on espère approcher au mieux la réalité. C'est une approche dite « fréquentiste ». La mesure est reine.

Cependant, le nombre de répétitions n'est pas toujours très élevé. De plus, on a parfois d'autres informations sur les valeurs attendues qu'il faut prendre en compte. Cela peut être la forme de la loi de distribution ou la certitude que l'intégrale nette doit être positive, comme dans le chapitre précédent. Comment inclure ces informations dans l'analyse des données ? C'est l'objet de l'approche dite bayésienne.

En ce qui concerne l'évaluation de type B de l'incertitude, on essaye aussi de définir l'incertitude à partir des quelques informations en notre possession. La probabilité représente alors la plausibilité de la valeur, pas sa fréquence par tirage.

Le GUM a donc une démarche qui n'est pas homogène : l'approche est fréquentiste pour l'évaluation de type A et « plausibiliste » pour celle de type B. Dans la littérature de métrologie et les normes les plus récentes, on peut noter une tendance à tout traiter avec une approche bayésienne. Ce chapitre a pour but d'introduire ces méthodes, sachant qu'il existe des livres complets sur le sujet.

Je me suis servi de deux articles pédagogiques très lumineux. Il s'agit de *Measurement as Inference : Fundamental Ideas* de W. Tyler Estler [8] et de *Where do we stand on maximum entropy?* de E.T. Jaynes [9]. Le deuxième est disponible gratuitement en ligne. Je mentionnerai quelques articles historiques sur le développement des concepts au cours de ce chapitre. Un des premiers articles (le premier ?) « militant » pour une approche bayésienne du calcul d'incertitude date de 1992 [17].

6.1 Un peu de théorie des probabilités

6.1.1 Un peu d'algèbre

Nous avons déjà introduit quelles notions sur les probabilités dans le chapitre 4. Nous avons vu que

$$p(A \cap B) = p(A) \cdot p(B) \quad (6.1)$$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B). \quad (6.2)$$

Mais **attention**, la relation (6.1) n'est vraie que si A et B sont indépendants. En effet, supposons que A signifie « tirer un cœur dans un jeu de 52 cartes » et B « tirer une carte rouge dans ce même jeu ». Comme les cœurs sont rouges, la relation (6.1) n'est plus correcte. Pour pouvoir traiter les cas où A et B sont corrélés, on va introduire les probabilités conditionnelles dans la section suivante.

A et B sont dits disjoints si $p(A \cap B) = 0$. C'est le cas par exemple si A signifie « tirer un cœur dans un jeu de 52 cartes » et B « tirer une carte noire dans ce même jeu ». Dans ce cas, l'équation (6.2) devient $p(A \cup B) = p(A) + p(B)$.

En notant \bar{A} , la proposition inverse de A , on a $p(A \cap \bar{A}) = 0$ et aussi $p(A) + p(\bar{A}) = 1$.

Les exemples présentés ici avaient tous une interprétation fréquentiste. Mais Cox [14] a démontré que ces règles étaient générales et uniques pour toute probabilité. Même s'il n'y a pas d'interprétation fréquentiste comme pour la probabilité de trouver de la vie sur Mars par exemple.

6.1.2 Probabilité conditionnelle

Quand nous avons raisonné sur les dés, nous avons fait l'hypothèse que les dés utilisés n'étaient pas pipés. Les résultats déduits dépendaient de cette hypothèse. En toute rigueur, il faudrait noter $p(6|H)$ la probabilité de tirer un 6 en supposant vraie l'hypothèse H qui est « les dés ne sont pas pipés ». La démarche bayésienne va consister « peser » cette hypothèse par des expériences.

En tenant compte explicitement des hypothèses dans la formulation des probabilités, les règles rappelées à la section précédente deviennent,

$$p(A \cap B|H) = p(A|H) \cdot p(B|AH) = p(B|H) \cdot p(A|BH) \quad (6.3)$$

$$p(A \cup B|H) = p(A|H) + p(B|H) - p(A \cap B|H). \quad (6.4)$$

Reprenons le cas d'un jeu de 52 cartes pour comprendre l'équation (6.3). L'hypothèse H est « le jeu n'est pas truqué ». On peut calculer de deux manières différentes la probabilité de tirer un as (A) rouge (R), $p(A \cap R|H)$. C'est la probabilité de tirer un as $p(A|H) = 4/52$, multipliée par la probabilité que cet as soit rouge $p(R|AH) = 2/4$. Ou bien, c'est la probabilité de tirer une carte rouge $p(R|H) = 1/2$ multipliée par la probabilité d'avoir un as parmi les cartes rouges $p(A|RH) = 2/26$. Les deux chemins mènent bien au même résultat.

L'intérêt de cette formulation est que l'on peut prendre en compte les corrélations éventuelles. Ainsi, la probabilité de tirer un cœur (C) rouge (R) est

$$p(C \cap R|H) = p(C|H) \cdot p(R|CH) = \frac{1}{4} \cdot 1 = \frac{1}{4}, \quad (6.5)$$

ou bien

$$p(C \cap R|H) = p(R|H) \cdot p(C|RH) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}. \quad (6.6)$$

Supposons maintenant qu'il y ait plusieurs hypothèses B_i possibles, qui soient disjointes et complètes, c'est à dire, telles que $p(B_i \cap B_j) = 0$ et $\sum p(B_i) = 1$. Avec le jeu de 52 cartes, ce peut être par exemple, cœur, carreau, pic et trèfle. Ce peut être aussi B et \bar{B} . On aura alors

$$p(A|H) = \sum_i p(A \cap B_i|H) = \sum_i p(A|B_iH) \cdot p(B_i|H). \quad (6.7)$$

Cette somme peut parfois être une intégrale. En reprenant l'exemple du jeu de cartes, cette expression peut se traduire par la probabilité de tirer un as est égale à la probabilité de tirer un as parmi les cœurs multipliée par la probabilité de tirer un cœur, plus celle de tirer un as parmi les carreaux multipliée par la probabilité de tirer un carreau, plus la même chose pour le pique et le trèfle, avec l'hypothèse H que le jeu n'est pas truqué. Cette équation peut aussi être vue comme la projection sur une base orthogonale complète, comme en mécanique quantique où les fonctions d'onde représentent aussi des densités de probabilité.

Evidemment, avec le jeu de cartes où le nombre de cartes par couleur est le même, cela ne présente pas beaucoup d'intérêt. Prenons donc une autre exemple. Supposons qu'un piège puisse attraper aussi bien des souris et des musaraignes, sans distinction et qu'il y a 60% de musaraignes. S'il y a 57% de femelles chez les souris et 53% chez les musaraignes, combien de chances a-t-on de capturer une femelle ?

Pour écrire cela sous forme mathématique, on note $p(S|H)$ la probabilité d'attraper une souris et $p(M|H)$ celle d'attraper une musaraigne avec l'hypothèse H que le piège ne fait pas la distinction. La proportion de femelles chez les souris est $p(F|S)$ et celle chez les musaraignes $p(F|M)$. La probabilité d'attraper une femelle est donc, en appliquant la formule (6.7),

$$p(F|H) = p(F|M) \cdot p(M|H) + p(F|S) \cdot p(S|H) = 0,53 \times 0,60 + 0,57 \times 0,40 = 54,6\%. \quad (6.8)$$

Rappelez-vous la citation de Laplace mise en exergue au début de ce chapitre...

6.1.3 Théorème de Bayes

Le théorème de Bayes, parfois appelé Bayes-Laplace, n'est qu'une réécriture de l'équation (6.3),

$$p(B|AH) = \frac{p(B|H) \cdot p(A|BH)}{p(A|H)}. \quad (6.9)$$

Il permet d'inverser les probabilités, c'est à dire, de calculer la probabilité de B si A , connaissant la probabilité de A si B . Par exemple, si l'on connaît la probabilité qu'un test médical donne un résultat positif si le patient est malade, on pourra calculer la probabilité que le patient soit malade si le test est positif.

Au chapitre précédent, nous avons aussi implicitement inversé des probabilités : nous étions parti d'une hypothèse sur la valeur vraie du mesurande et nous en avons déduit des informations sur la distribution de probabilité des valeurs prises par le mesurande. Puis, nous avons inversé le processus pour en déduire des informations sur la probabilité d'obtenir une valeur vraie donnée à partir d'une valeur mesurée ou d'une moyenne de valeurs mesurées.

Ecrit sous cette forme, ce résultat est dû à Pierre Simon, Marquis de Laplace, ancien étudiant de l'université de Caen, dans son traité philosophique des probabilités [11], même s'il est connu sous le nom de théorème de Bayes. T. Bayes avait énoncé un résultat similaire dans un cas particulier quelques années auparavant [10].

En écrivant l'expression (6.9) pour un B_j particulier, et en utilisant l'équation (6.7) pour le dénominateur, on a finalement,

$$p(B_j|AH) = \frac{p(B_j|H) \cdot p(A|B_jH)}{\sum_i p(A|B_iH) \cdot p(B_i|H)}. \quad (6.10)$$

Il s'agit d'une autre forme du théorème de Bayes est aussi due à Laplace [11]. Dans le cadre de l'application aux calculs d'incertitude qui nous intéressent ici, la somme du dénominateur, qui correspond à une normalisation, peut parfois prendre la forme d'une intégrale.

6.1.4 Inférence bayésienne

On a déjà dit que dans ce chapitre, il fallait abandonner l'interprétation fréquentiste des probabilités pour leur préférer l'interprétation plausibiliste. De façon caricaturale, l'approche fréquentiste correspond à la logique shadok du dessin de la figure 6.1. Comment prendre en compte le fait que les échecs multiples changent la plausibilité et donc la probabilité de réussite ?

C'est le théorème de Bayes qui permet de répondre à cette question : l'équation (6.9) donne l'expression de la probabilité que B soit vraie sachant A et H à partir de la probabilité que B soit vraie sachant H seulement. $p(B|AH)$ prend en compte le retour d'expérience A .

Prenons un exemple simple pour comprendre comment cela marche. Avant de proposer un traitement lourd à un malade, un médecin va faire procéder à un examen médical. La proposition B est donc « le patient a la maladie X ». Après un examen sanguin, le patient a un résultat positif ou négatif. La proposition A est « le test est positif ». Mais le test n'est pas fiable à 100% : il a une probabilité $p(A|B)$ d'être positif si le patient est malade et une probabilité $p(A|\bar{B})$ si le patient n'est pas malade. Si le résultat du test sanguin est positif, quelle est la probabilité que la patient soit malade ?

Le théorème de Bayes donne,

$$p(B|A) = \frac{p(B) \cdot p(A|B)}{p(A)}. \quad (6.11)$$

$p(B)$ est la probabilité que le patient ait la maladie X estimée par le médecin avant le test. On l'appelle le **prior** [prior]. On peut imaginer plusieurs cas :

- c'est un simple dépistage et le médecin n'avait pas d'a priori. On peut donc attribuer $1/2$ à $p(B)$;
- le médecin soupçonne déjà la maladie X à partir d'autres indications cliniques. Il pourra attribuer à $p(B)$ une valeur plus proche de 1 ;
- la maladie X est rare, mais le médecin fait faire le test par conscience professionnelle. La valeur attribuée à $p(B)$ pourra alors être inférieure à $1/2$.

$p(A|B)$ est donné par le fabricant du test et vient d'un tude statistique. Ne reste plus qu'à connaître le dénominateur $p(A)$. Comme il y a plusieurs façons d'avoir un test positif, on est obligé de prendre en compte

Les devises Shadok



EN ESSAYANT CONTINUUELLEMENT
ON FINIT PAR RÉUSSIR. DONC :
PLUS ÇA RATE, PLUS ON A
DE CHANCES QUE ÇA MARCHE.

FIG. 6.1 – Logique shadok. (dessin trouvé sur Internet)

tous les cas en utilisant la deuxième forme du théorème de Bayes. Ainsi,

$$p(B|A) = \frac{p(B) \cdot p(A|B)}{p(A|B)p(B) + p(A|\bar{B})p(\bar{B})} = \frac{p(B) \cdot p(A|B)}{p(A|B)p(B) + p(A|\bar{B})(1 - p(B))} \quad (6.12)$$

Application numérique. On va distinguer plusieurs cas :

- Le test est fiable à 100% : $p(A|B) = 1$ et $p(A|\bar{B}) = 0$. Alors $p(B|A) = 1$, quel que soit le prior. Il y a donc une probabilité de 100% que le patient soit malade si le test est positif.
- On a $p(A|B) = 0,90$ et $p(A|\bar{B}) = 0,05$. Si $p(B) = 1/2$, alors $p(B|A) = 0,90/0,95 = 0,95$. Le test a bien fait progresser la connaissance !
- On a toujours $p(A|B) = 0,90$ et $p(A|\bar{B}) = 0,05$. Si $p(B) = 0,80$, alors $p(B|A) = 0,99$!
- On peut imaginer que le médecin est sûr de son diagnostic avec $p(B) = 0,80$, mais que les tests ne sont pas très fiables : $p(A|B) = 0,80$ et $p(A|\bar{B}) = 0,2$. Alors $p(B|A) = 0,9$. Le test n'est pas très utile...

On va maintenant appliquer le théorème de Bayes à la mesure. La question qui se pose donc à toute inférence bayésienne, est la détermination du prior. Il existe des règles en métrologie. C'est l'objet de la section suivante.

6.2 Maximisation de l'entropie de l'information

Pour une information donnée, chaque personne doit proposer le même prior. Pour cela, il faut des règles.

Supposons que l'on joue à pile ou face et que la pièce est biaisée : une des deux faces a plus de chance de sortir que l'autre, mais on ne sait pas laquelle. Alors, on n'a pas d'autre choix que de supposer initialement que la probabilité est de $\frac{1}{2}$ pour chacune des faces. Après plusieurs lancers, cette probabilité pourra être réévaluée à l'aide du théorème de Bayes.

C'est la même idée qui sous-tend le choix du prior pour l'évaluation de l'incertitude. La question qui se pose est comment exprimer notre ignorance ? Les deux postulats sur lesquels repose la démarche sont :

1. Les probabilités représentent l'état de nos connaissances. Elles ne mesurent pas forcément la fréquence d'un évènement, mais plutôt sa plausibilité.
2. Si l'on n'a pas de raison de penser que A est plus probable que B , alors, on écrit $p(A) = p(B)$. C'est le **principe de la raison insuffisante** [principle of insufficient reason] de Jacob Bernoulli. En 1921 Keynes proposa une formulation plus pertinente : **principe d'indifférence** [principle of indifference]

Si l'on sait uniquement, par exemple, que les valeurs pouvant être prises par une grandeur donnée appartiennent à l'intervalle $[a, b]$, quel prior choisir ? Si l'on divise cet intervalle en n sous-intervalles de longueur identique, alors comme pour la pièce biaisée, on va affecter la même probabilité à chaque sous intervalle. Comme le résultat ne doit pas dépendre du nombre de sous-intervalles, la seule possibilité est de prendre une probabilité constante sur l'intervalle $[a, b]$. C'est ce que nous avons fait pour l'évaluation de type B de l'incertitude.

Cette approche est difficilement applicable dans des situations beaucoup plus compliquées. On utilise maintenant le concept d'entropie de l'information introduit par Shannon [15] que l'on va chercher à maximaliser en tenant compte des informations solides disponibles.

Cette entropie est définie par

$$S = - \sum_i p_i \ln(p_i) \quad (6.13)$$

dans les cas discrets et

$$S = - \int p(x) \ln(p(x)) dx \quad (6.14)$$

dans les cas continus.

Reprenons le cas de la pièce biaisée pour laquelle il n'y a que deux possibilités.

$$S = -p_1 \ln(p_1) - p_2 \ln(p_2) = -p_1 \ln(p_1) - (1 - p_1) \ln(1 - p_1). \quad (6.15)$$

En dérivant par p_1 pour trouver le maximum, on obtient $p_1 = 1/2$.

De la même façon, si l'on sait uniquement que les valeurs pouvant être prises par une grandeur donnée appartiennent à l'intervalle $[a, b]$, la maximisation de l'entropie de l'information conduit à une distribution constante sur cet intervalle. Je vous laisse le vérifier à titre d'exercice. Je vais traiter un autre cas classique pour montrer la méthode.

Exemple : Soit $p(x)$ une distribution à une dimension dont on connaît la valeur moyenne μ et la variance σ^2 . On va la déterminer en maximisant l'entropie

$$S = - \int p(x) \ln(p(x)) dx, \quad (6.16)$$

en tenant compte des contraintes

$$\mu = \int x p(x) dx, \quad \sigma^2 = \int x^2 p(x) dx \quad \text{et} \quad 1 = \int p(x) dx, \quad (6.17)$$

à l'aide des multiplicateurs de Lagrange α , β et γ . Cela conduit à trouver la fonction $p(x)$ qui maximise l'intégrale,

$$I[p(x)] = \int p(x) [-\ln(p(x)) + \alpha x + \beta x^2 + \gamma] dx. \quad (6.18)$$

Si la solution $p_0(x)$ maximise $I[p(x)]$, alors $I[p_0(x) + \delta p(x)]$ sera égal à $I[p_0(x)]$ au premier ordre en $\delta p(x)$. En faisant un développement au premier ordre, on obtient

$$I[p_0 + \delta p] = \int (p_0 + \delta p) [-\ln(p_0 + \delta p) + \alpha x + \beta x^2 + \gamma] dx \quad (6.19)$$

$$= \int p_0(x) (-\ln(p_0(x)) + \alpha x + \beta x^2 + \gamma) dx + \int \delta p(x) [-1 - \ln(p_0(x)) + \alpha x + \beta x^2 + \gamma] dx \quad (6.20)$$

Le deuxième terme de l'équation précédente doit être nul quel que soit $\delta p(x)$ et donc,

$$-1 - \ln[p_0(x)] + \alpha x + \beta x^2 + \gamma. \quad (6.21)$$

$p_0(x)$ est donc de la forme,

$$p_0(x) = \exp[\gamma - 1 + \alpha x + \beta x^2]. \quad (6.22)$$

En ajustant les multiplicateurs de façon à satisfaire aux contraintes, équations (6.17), on a finalement,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2}\right] \quad (6.23)$$

qui est une belle gaussienne! Ainsi, la distribution qui maximise l'entropie de l'information avec comme condition, une valeur moyenne et une variance données, est une gaussienne.

Dans cet exemple, on a vu comment traiter mathématiquement notre ignorance à travers l'entropie et nos connaissances à travers les multiplicateurs de Lagrange pour en déduire la distribution qui prend tout en compte.

Les nouvelles normes internationales recommandent d'utiliser le principe de maximisation de l'entropie de l'information pour l'estimation de type B de l'incertitude et pour le prior des méthodes bayésiennes. Le supplément 1 au GUM [4] donne une liste de nombreuses distributions qui maximisent l'entropie de l'information en fonction de contraintes spécifiques et explique comment faire un tirage aléatoire.

6.3 Application aux incertitudes

6.3.1 Retour sur les seuils de décision et limites de détection

Dans le chapitre précédent, le calcul des seuils de décision et des limites de détection a été fait à partir de méthodes dites bayésiennes. Où intervenait le théorème de Bayes? Reprenons la démonstration.

Supposons, comme dans le chapitre précédent, que le nombre de coups détectés satisfait à une distribution de Poisson. Connaissant la valeur vraie de ce nombre, μ , on a la probabilité de détecter N coups qui est égale à

$$P(N|\mu) = \frac{e^{-\mu} \mu^N}{N!}. \quad (6.24)$$

Mais, ce qui nous intéresse, c'est $P(\mu|N_1)$, la distribution de probabilité de la valeur vraie μ connaissant un nombre de coups détectés N_1 . L'inversion se fait à l'aide du théorème de Bayes :

$$P(\mu|N_1) = \frac{P(N_1|\mu)f(\mu)}{P(N_1)}. \quad (6.25)$$

Ici, $P(N_1|\mu)$ est la distribution de Poisson et $f(\mu)$, le prior. En l'absence d'information particulière, le prior est pris constant sur l'intervalle $[0, +\infty[$. Le dénominateur doit prendre en compte les valeurs de N_1 dues à tous les μ possibles,

$$P(N_1) = \int_0^{+\infty} P(N_1|\mu) f(\mu) d\mu. \quad (6.26)$$

Le calcul de l'intégrale se fait aisément en intégrant N_1 fois par partie et donne $f(\mu)$.

On obtient finalement,

$$P(\mu|N_1) = \frac{e^{-\mu} \mu^{N_1}}{N_1!}. \quad (6.27)$$

Bien que cette distribution ressemble beaucoup à une distribution de Poisson, ce n'en est pas une. Ici, N_1 est considéré comme un paramètre constant et μ comme une variable aléatoire. C'est l'inverse pour la distribution de Poisson.

On peut calculer aisément sa valeur moyenne et sa variance. L'espérance

$$E(\mu|N_1) = \int_0^{+\infty} \mu P(\mu|N_1) d\mu = (N_1 + 1) \int_0^{+\infty} \frac{e^{-\mu} \mu^{N_1+1}}{(N_1 + 1)!} = N_1 + 1. \quad (6.28)$$

On montre de même que la variance vaut aussi $N_1 + 1$. Cela justifie a posteriori la règle de remplacer les taux de comptages n par $n + 1$ dans le chapitre précédent.

Pour bien comprendre la démarche adoptée, il faut garder en tête que le résultat d'un mesurage (N_1 ici) n'est plus considéré comme une variable aléatoire dans l'approche bayésienne, mais comme une donnée. A partir de cette donnée, on cherche la probabilité d'obtenir la valeur vraie μ . Ici, la probabilité doit bien être considérée comme la plausibilité d'avoir une valeur donnée de μ . Et à partir de cette probabilité, on peut définir le seuil de décision et la limite de détection.

Pour aller plus loin, voir par exemple l'article de Vivier et al [21].

6.3.2 Evaluation de type A de l'incertitude

On l'a déjà dit, l'évaluation de type A de l'incertitude est basée sur une approche fréquentiste. Il existe dans la littérature scientifique une approche bayésienne qui change l'expression et même des articles qui appellent à une révision du GUM en ce sens. Voir par exemple celui de Kacker et Jones [22].

Soient n résultats d'un même mesurage, z_1, \dots, z_n caractérisés par leur valeur moyenne \bar{z} et sa variance $s(\bar{z})$. Ces données sont fixes et l'incertitude est dans notre connaissance de la valeur vraie de la grandeur Z , objet du mesurage.

En supposant que l'on pas d'a priori sur les valeurs prises par Z , on va choisir un prior non informatif. Je ne ferai pas le calcul, car il faut inverser la distribution de Student l'aide du théorème de Bayes et c'est compliqué. La distribution de probabilité pour la valeur vraie est caractérisée par une espérance

$$E(Z) = \bar{z}, \tag{6.29}$$

et un écart-type

$$u_B(\bar{z}) = \sqrt{\frac{n-1}{n-3}} s(\bar{z}). \tag{6.30}$$

Pour n suffisamment élevé, le terme correctif ne change rien et les approches fréquentiste et bayésienne coïncident. Ce n'est plus le cas, pour n faible qui impose que $n > 3$ pour être valable.

Comment faire si ce n'est pas le cas? Il faut faire des hypothèses sur le prior de façon à avoir une autre expression. C'était le cas précédemment où l'on savait que la radioactivité est bien décrite par la statistique de Poisson. Ce n'est pas la peine de multiplier les comptages pour le vérifier à chaque fois. Avec un seul comptage, on connaît la variance et l'on peut utiliser cette connaissance à l'aide du théorème de Bayes.

Annexe A

Historique de la normalisation du calcul d'incertitude

Ce texte est tiré du site de l'AFNOR, mais il n'est plus accessible en ligne actuellement.

« En 1980, en une page, le Comité international des poids et mesures (CIPM) avait publié sa recommandation INC-1 (1980) sur la classification en deux catégories des composantes de l'incertitude d'un résultat de mesure (en fonction de leur évaluation par des méthodes statistiques ou par d'autres moyens), sur son expression sous la forme d'écart types et sur un éventuel facteur d'élargissement.

Une deuxième recommandation 1 (CI-1981) fait en 1981 la promotion de la première recommandation pour en étudier la mise en application sur deux à trois ans en vue d'une uniformisation de l'expression des incertitudes de mesure.

Une troisième recommandation 1 (CI-1986) transforme en 1986 la recommandation de 1980 en obligation pour tous les organismes travaillant dans le cadre du CIPM.

Entre temps, l'ISO/TAG 4 avait créé son deuxième groupe de travail pour publier un document expliquant la recommandation de 1980 et donnant des conseils quant à sa mise en application. Ce groupe de travail, comme celui du VIM, rassemblait les représentants des sept organismes internationaux directement concernés par la métrologie.

Le résultat de ce groupe de travail fut la publication en 1995 du Guide pour l'expression de l'incertitude de mesure, appelé "GUM" dans notre jargon. Nous passions en quinze ans d'une page réservée aux spécialistes à 105 pages déjà plus "démocratisées", mais encore caractérisées comme rébarbatives (ce délai très long ne doit pas être critiqué compte tenu du sujet, des cultures et des intérêts mis en jeu, dont nous constatons encore les effets aujourd'hui).

Le GUM fut repris par le CEN sous la référence ENV 13005 (l'ENV étant équivalente à notre statut de norme expérimentale, ce qui n'est pas très valorisant ni fédérateur pour ce genre de document). Cela ne se fit pas sans polémiques, ce qui explique ce statut, les polémiques étant liées au mode d'élaboration du document (ce n'est pas une norme internationale rédigée comme telle, même s'il en a la valeur et qu'il soit cité comme document de référence pour l'élaboration des normes dans les directives de l'ISO/CEI et le règlement du CEN/CENELEC), à l'existence d'autres documents normatifs sur les incertitudes de mesure soi-disant contradictoires (la série des normes ISO 5725) et à l'absence de structure horizontale pour la métrologie au niveau européen.

Cette ENV 13005 fut ensuite reprise dans la collection nationale, mais avec le statut de norme homologuée.

Bien que le GUM soit devenu la référence puisqu'il était dans la lignée de la recommandation de 1980 du CIPM, il lui fut reproché d'avoir été rédigé par des métrologues pour l'application aux étalonnages en laboratoire, contexte bien particulier où toutes les conditions d'environnement et autres facteurs d'influence sont censés être identifiés et maîtrisés.

À la même époque, le domaine des essais avait aussi pris conscience de l'importance de cette notion d'incertitude de mesure associée au résultat de mesure d'un essai, pour pouvoir faire des comparaisons, estimer les risques et prendre des décisions en connaissance de cause. Mais la maîtrise des conditions d'environnement et des facteurs d'influence d'un essai n'est pas forcément du même niveau que celui d'un étalonnage, et c'est le Sous-comité 6 "Méthodes et résultats de mesure" du Comité technique n 69 de l'ISO "Application des méthodes statistiques" qui apporta des solutions avec la première norme ISO 5725 publiée en décembre 1987, cette norme

traitant de l'exactitude (justesse et fidélité) des résultats et méthodes de mesure, des questions de répétabilité et de reproductibilité des méthodes de mesure normalisées et des comparaisons interlaboratoires.

De là à créer deux chapelles (une troisième pouvant être constituée de ceux voulant ignorer dans leur secteur ces questions d'incertitude de mesure, ne se sentant pas concernés et ne voulant surtout rien faire ni rien changer dans leurs pratiques) et à traiter l'autre d'hérétique, le pas fut vite franchi.

Nos experts gardèrent la tête froide (heureusement pour ses travaux sur les incertitudes de mesure, notre Commission générale "Métrologie dans l'entreprise" avait eu la sagesse de se rapprocher de la commission française en charge des méthodes statistiques et "miroir" de l'ISO/TC 69, et de constituer des groupes de travail composés de métrologues et de statisticiens)! Ils travaillèrent à la rédaction d'un fascicule de documentation fédérateur, expliquant qu'entre l'expression du besoin d'incertitude de mesure et le résultat de l'incertitude de mesure il y avait deux voies parallèles suivant le contexte (celui des essais ou celui de l'étalonnage). Le GUM et les normes ISO 5725 deviennent alors complémentaires d'une certaine manière. Le traitement dans le fascicule de documentation d'une dizaine d'exemples dans différents secteurs d'activité confirme que les notions d'incertitudes de mesure sont générales.

Ce fascicule de documentation a été publié en 1999 sous la référence FD X07-021, Normes fondamentales Métrologie et applications de la statistique Aide à la démarche pour l'estimation et l'utilisation de l'incertitude des mesures et des résultats d'essais. Une version anglaise de ce fascicule de documentation a été remise à l'ISO/TC 69 comme contribution à ses travaux.

En 2000, il a été envisagé de réviser le GUM, question qui se pose systématiquement environ tous les trois ans pour tout document normatif. De la réflexion sur cette révision, il est ressorti qu'il ne fallait pas toucher au GUM dans l'immédiat, mais seulement le compléter par la publication de documents sur, par exemple, la méthode de simulation de Monte Carlo, l'utilisation de l'incertitude, l'interpolation, etc.

Un premier supplément au GUM, le Supplément 1 : Méthodes numériques pour la propagation de distributions, est en cours d'élaboration. »

Bibliographie

Normes internationales

- [1] Bureau International des poids et mesures, *Le système international d'unités*, (2006), <http://www.bipm.org/>
- [2] *Vocabulaire international de métrologie - Concepts fondamentaux et généraux et termes associés* (2008), <http://www.bipm.org/>. Les principaux termes de métrologie y sont définis en français et en anglais.
- [3] *Guide pour l'expression de l'incertitude de mesure*, NF ENV 13005 (1999), <http://www.bipm.org/>.
- [4] *Evaluation des données de mesure - Supplément 1 du "Guide pour l'expression de l'incertitude de mesure" - Propagation de distributions par une méthode de Monte Carlo* (2008) <http://www.bipm.org/>.
- [5] *Détermination des limites caractéristiques (seuil de décision, limite de détection et extrémités de l'intervalle de confiance) pour mesurages de rayonnements ionisants - Principes fondamentaux et applications*, NF ISO 11929. Une version anglaise est aussi disponible. La version internationale, en anglais, date de 2009 et la version française de 2010.

Livres

- [6] John R. Taylor, *An introduction to error analysis, the study of uncertainties in physical environments*, University Science Books, (1982). Il existe une version française publiée chez Dunod (2000).

Articles

Articles pédagogiques

- [7] A. Vivier, R. Fottorino et B. Rousse, *Seuil de décision et limite de détection : estimation, interprétation et optimisation, première partie : les principes de base*, Radioprotection **45** (2010) 321
- [8] W. Tyler Estler, *Measurement as Inference : Fundamental Ideas*, Annals of the CIRP **48** (1999) 611
- [9] E.T. Jaynes, *Where do we stand on maximum entropy?*, in The Maximum Entropy Formalism, R. D. Levine and M. Tribus (eds.), M. I. T. Press, Cambridge, MA (1979) 15. <http://bayes.wustl.edu/etj/articles/stand.on.entropy.pdf> Article fascinant!

Articles historiques

- [10] T. Bayes, *An essay towards solving a Problem in the Doctrine of Chances*, Phil. Trans. Roy. Soc. (1763) 330. Repris dans Biometrika **45** (1958) 293 et ici : <http://rstl.royalsocietypublishing.org/content/53/370.full.pdf>. Ou encore ici avec une écriture plus moderne : <http://www.stat.ucla.edu/history/essay.pdf>. La lecture de ce texte est assez difficile.
- [11] P.S. de Laplace, *Essai philosophique sur les probabilités* (1774). Voir aussi *Théorie analytique des probabilités*, imprimerie royale (1847), réédité par Gabay (1995). Textes consultables en ligne sur le site de la BNF : <http://gallica.bnf.fr>. Le premier texte, qui ne contient aucune équation, est très lumineux.

- [12] B.L. Welch, *The significance of the difference between two means when the population variances are unequal*, *Biometrika* **29** (1937) 350
- [13] F.E. Satterthwaite, *An approximate distribution of estimates of variance components*, *Biometrics Bulletin* **2** (1946) 110
- [14] R.T. Cox, *Probability, Frequency and Reasonable Expectation*, *Am. J. Phys.* **14** (1946) 1
- [15] Claude E. Shannon, *A Mathematical Theory of Communication*, *Bell System Tech. J.* **27** (1948) 379 et 623. Disponible en ligne.
- [16] E.T. Jaynes, *Information Theory and Statistical Mechanics*, *Phys. Rev.* **106** (1957) 620
- [17] K. Weise et W. Wöger, *A Bayesian theory of measurement uncertainty*, *Meas. Sci. Technol.* **3** (1992) 1

Autres articles

- [18] Donald W. Marquardt, *An Algorithm for Least-Squares Estimation of Nonlinear Parameters*, *J. Soc. Industr. Appl. Math.* **11** (1963) 431. Disponible en ligne gratuitement.
- [19] W. Zimmermann, *Evaluation of Photopeaks in Scintillations Gamma-Ray Spectroscopy*, *Rev. Sci. Instr.* **32** (1961) 1063
- [20] T. Mukoyama, *Fitting of Gaussian to peaks by non-iterative method*, *Nucl. Instr. Meth.* **125** (1975) 289
- [21] A. Vivier and J. Aupiais, *Optimization of the decision threshold for single radioactive counting*, *Radiochim. Acta* **95** (2007) 477
- [22] R. Kacker and A. Jones, *Metrologia* **40** (2003) 235