

# **Analyse quantitative**

CHRISTOPHE ROUSSEL  
[Année 2019-2020]

# Analyse quantitative : approche statistique

## I Introduction

L'analyse quantitative requiert le traitement des données analytiques par une approche statistique. La statistique est la science qui a pour objet de recueillir un ensemble de données numériques relatives à tel ou tel phénomène aléatoire et d'exploiter cette information pour établir toutes les relations de causalité par l'analyse pour son interprétation.

Un **phénomène aléatoire** est un phénomène comportant des variables aléatoires, c'est à dire des variables liées au hasard et dont les valeurs ne peuvent, en conséquence, être connues d'avance comme par exemple le nombre de points marqués par un dé lorsqu'on le lance.

Dans le concept de la statistique, on distingue :

**La statistique descriptive**, ou statistique de constatation, qui concerne les tableaux de données, tableaux relatifs à des inventaires, les graphiques, des recensements...

**La méthode statistique** qui concerne l'ensemble des procédés et méthodes utilisées pour l'analyse et l'interprétation des données.

### I.1 Domaine d'application

Le domaine d'utilisation de la statistique est tellement étendu qu'il ne nous est impossible de citer toutes ses applications. Nous ne donnerons ici seulement que quelques exemples tels que : les domaines de recherche biologique, médicale, spatiale ; le contrôle de fabrication dans l'industrie ; les sondages d'opinion, enquêtes de marché, assurances ; la recherche opérationnelle ; les études de conjoncture, économiques...

Dans tous les cas de figure, on doit définir un ensemble, ou référentiel statistique, composé d'éléments ou unités statistiques est appelé "population" ou "univers".

## I.2 Caractères qualitatifs et quantitatifs, continus et discrets

### I.2.1 Les caractères qualitatifs et quantitatifs

Les caractères **quantitatifs** sont ceux auxquels on peut attribuer une valeur numérique. En revanche, les caractères **qualitatifs** sont ceux auxquels on ne peut seulement associer qu'une valeur arbitraire, une indexation. Par exemple une **taille** est un caractère **quantitatif** alors qu'une **couleur** est un caractère **qualitatif**.

Un ensemble ordonné de valeurs de caractère **quantitatif** constitue une suite ou **série statistique**. Un ensemble ordonné conventionnellement (indexation) de caractère **qualitatif** constitue une **nomenclature**.

Dans le cadre de la convention fixée pour l'indexation (règle normalisée ou prescription légale), la nomenclature est parfois désignée sous le nom de code comme par exemple le code des départements applicables à l'immatriculation des automobiles françaises.

### I.2.2 Les caractères continus et discrets

Un caractère **continu** est un caractère qui peut prendre n'importe quelle valeur numérique comme par exemple une masse, une surface, un prix...

Un caractère discret (ou discontinu) est un caractère qui ne peut prendre que certaines valeurs (en général, des nombres entiers) comme par exemple le nombre de personnes constituant une famille.

Dans le cas d'un caractère discret, l'interpolation est dénuée de sens. Très souvent on se sert de la représentation graphique pour illustrer les résultats. Les histogrammes et les distributions s'avèrent être des outils performants, clairs dans l'énonciation et compréhensifs vis-à-vis du lecteur. Les perfectionnements considérables des outils informatiques ont contribués à étendre les possibilités de la statistique. De nos jours, avec les nombreux logiciels à disposition, la représentation graphique des résultats ne pose plus guère de problèmes et l'avantage est évident; l'appréciation instantanée de la qualité du travail en tirant les conclusions qui s'imposent. Par exemple, une courbe symétrique provoque auprès de l'expérimentateur une satisfaction personnelle non cachée; qu'elle soit justifiée, reste encore à prouver. Les distributions à plusieurs sommets sont à considérer avec tout le respect que cela demande, il faut méditer sur le résultat et s'intéresser à d'éventuelles anomalies qui font surface.

## II Les sources d'erreur et leurs quantifications

### II.1 Introduction

Les sources d'erreur qui entachent la précision d'un résultat analytique sont de trois sortes :

$$Qualité\ d'analyse \propto \begin{cases} justesse\ (erreurs\ systématiques) \\ précision\ (erreurs\ aléatoire) \\ sensibilité\ de\ la\ méthode\ analytique \end{cases}$$

Nous allons à présent, à travers des exemples, s'intéresser à ces différentes sources d'erreur ainsi qu'à leur quantification afin d'analyser aux mieux un résultat analytique. Chaque fois que l'on fait une mesure, on exprime par une valeur numérique  $x_i$  l'estimation d'une grandeur, dont la valeur réelle, souvent inconnue, est  $x_0$ . La détermination est d'autant plus juste que la différence  $x_i - x_0$  est petite. Dans la plupart des cas, on effectue plusieurs mesures et l'on définit une moyenne. Si la moyenne la grandeur  $x$  est calculée à partir de  $N$  valeurs de  $x$  tel que  $N \leq 20$  alors elle sera notée :  $\bar{x}$ . Si par contre cette moyenne est calculée à partir de  $N$  valeurs de  $x$  tel que  $N > 20$ , alors elle sera notée :  $\mu_x$ . Quel que soit la valeur de  $N$ , on associe à la moyenne un écart-type qui se notera  $s_x$  pour  $N \leq 20$  et  $\sigma_x$  pour  $N > 20$ . Les calculs des moyennes et écart-types seront présentés ci-après. Dès lors, la valeur vraie est donc encadrée de la manière suivante :

$$\bar{x} - s_x \leq x_0 \leq \bar{x} + s_x \quad (N \leq 20) \quad \text{ou} \quad \mu_x - \sigma_x \leq x_0 \leq \mu_x + \sigma_x \quad (N > 20)$$

La qualité d'analyse est d'autant meilleure que l'intervalle autour de  $x_0$  est petit. Les différentes erreurs qui entachent la qualité d'analyse sont décrites ci-après.

### II.2 Moyenne

La moyenne arithmétique est la plus couramment utilisée. Pour le calcul de la moyenne, il est inutile de dresser la série dans un ordre croissant ou décroissant, les termes sont additionnés dans un ordre quelconque. Deux propriétés caractérisent la moyenne : (i) la somme algébrique des écarts de chaque valeur de la série par rapport à la moyenne est nulle ; (ii) la moyenne arithmétique est la grandeur pour laquelle la somme des carrés des écarts par rapport à la moyenne est minimum.

La moyenne arithmétique désignée par  $\bar{x}$  ( $N \leq 20$ ) ou  $\mu_x$  ( $N > 20$ ) se calcule de la manière suivante :

$$\bar{x} \text{ ou } \mu_x = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{\sum_{i=1}^{i=N} x_i}{N}$$

La moyenne arithmétique est une bonne approximation d'un ensemble d'une certaine population. Elle doit être la valeur la plus rapprochée de la moyenne idéale, souvent inconnue, qui représente la valeur vraie.

### II.3 Justesse : erreur systématique

La justesse ou encore erreur systématique en français est décrite par les termes "accuracy" en anglais et "systematische fehler" en allemand. Les erreurs systématiques sont caractérisées par un écart à la valeur vraie  $x_0$  soit positif, soit négatif, elles agissent toujours dans le même sens et affectent l'exactitude de la méthode. Cet écart à la valeur vrai peut être qualifié par rapport à une valeur précise d'un groupe de données ou par rapport à la moyenne du groupe de données. On peut donner comme exemple le mauvais calibrage d'une burette, un décalage de longueurs d'ondes d'un spectrophotomètre, un mauvais tarage de balance etc... Elles peuvent être corrigées lors de la détermination du facteur de décalibration. Si on peut quantifier cette erreur, par exemple au moyen d'une détermination à partir d'un étalon, le résultat doit être corrigé. Parfois, un facteur correctif est indiqué dans les méthodes d'analyse.

#### II.3.1 Ecart d'une valeur précise d'un groupe de donnée par rapport à la valeur vraie

Si on mesure une valeur  $x_i$  pour l'échantillon  $i$  dont la valeur réelle est  $x_0$  on peut alors donner l'erreur systématique absolue  $\Delta x_i$  ou relative  $E_r x_i$  (%) sur la mesure de l'échantillon  $i$  :

$$\Delta x_i(x_0) = x_i - x_0 \quad \text{et} \quad E_r x_i(x_0) = \frac{|\Delta x_i(x_0)|}{x_0} \times 100$$

### II.3.2 Ecart d'une moyenne d'un groupe de données par rapport à la valeur vraie

Dans le cas où la valeur trouvée pour la mesure de l'échantillon  $i$  est une moyenne déterminée à partir de  $N$  analyses :  $\bar{x}$  ( $N \leq 20$ ) ou  $\mu_x$  ( $N > 20$ ) on pourra la comparer avec la valeur vraie  $x_0$  :

$$\Delta\bar{x}(x_0) = \bar{x} - x_0 \quad \text{ou} \quad \Delta\mu_x(x_0) = \mu_x - x_0$$

$$E_r\bar{x}(x_0) = \frac{|\Delta\bar{x}(x_0)|}{x_0} \times 100 \quad \text{ou} \quad E_r\mu_x(x_0) = \frac{|\Delta\mu_x(x_0)|}{x_0} \times 100$$

### II.3.3 Ecart d'une valeur précise d'un groupe de donnée par rapport à moyenne du groupe de données

Dans bien des cas en chimie analytique, on ne connaît pas la valeur vraie  $x_0$  d'une mesure. Dans ce cas il faut approximer la valeur vraie  $x_0$  par la moyenne des valeurs obtenues sur l'échantillon  $i$  mesuré  $N$  fois ( $\bar{x}$  ( $N \leq 20$ ) ou  $\mu_x$  ( $N > 20$ )). Dès lors, on ne peut que comparer une mesure  $x_i$  par rapport à la moyenne  $\bar{x}$  :

$$\Delta x_i(\bar{x}) = x_i - \bar{x} \quad \text{ou} \quad \Delta x_i(\mu_x) = x_i - \mu_x$$

$$E_r x_i(\bar{x}) = \frac{|\Delta x_i(\bar{x})|}{\bar{x}} \times 100 \quad \text{ou} \quad E_r x_i(\mu_x) = \frac{|\Delta x_i(\mu_x)|}{\mu_x} \times 100$$

On peut calculer une erreur expérimentale moyenne  $\bar{d}$  sur  $N$  analyses comme :

$$\bar{d}_x = \frac{\sum_{i=1}^{i=N} |x_i - \bar{x}|}{N} \quad \text{ou} \quad \bar{d}_x = \frac{\sum_{i=1}^{i=N} |x_i - \mu_x|}{N}$$

Ici, on mesure en fait la dispersion de toutes les mesures par rapport à la moyenne des mesures. On visualise alors une erreur autour de la moyenne  $\bar{x}$  ou  $\mu_x$  qui caractérise la précision de la mesure. Cette approche simpliste doit être traitée d'un point de vue statistique, par le calcul de l'écart-type.

## II.4 Précision : erreur aléatoire

La précision ou encore erreur aléatoire en français est décrite par les termes "precision ou standard deviation" en anglais et "zufallsfehler" en allemand. La non reproductibilité des résultats est souvent due à un grand nombre de petites sources d'erreur qui peuvent jouer dans les deux sens, positif et négatif (erreurs aléatoires). Elles trouvent leur nature à la fois dans la méthode analytique et dans l'appareil de mesure lui-même. Elles affectent la précision de la mesure. On peut alors utiliser les résultats des mesures pour fixer la limite de confiance au moyen des méthodes statistiques.

### II.4.1 Définitions

**L'erreur aléatoire ou fortuite** se produit tantôt du côté positif, tantôt du côté négatif par rapport à la valeur moyenne. Pour un grand nombre de mesures, la somme algébrique de ses valeurs successives est nulle. Les erreurs aléatoires obéissent à la **loi de Laplace-Gauss (loi normale)**. Toute erreur est de par nature fonction de la méthode d'analyse choisie, des conditions de travail, de l'habileté de l'expérimentateur, de l'homogénéité du produit à analyser. La densité de probabilité  $f(x)$  de la loi normale est :

$$f(x) = \frac{A}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu_x}{\sigma_x} \right)^2}$$

où A est une constante de pondération et dont la représentation graphique est la suivante :

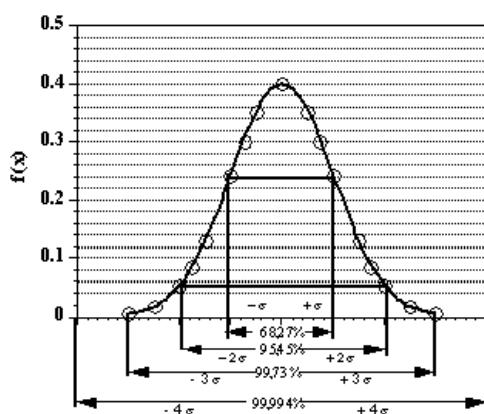


Figure 1. densité de probabilité de la loi normale

Nous allons à présent décrire les grandeurs statistiques en termes de calcul : moyenne, écart-type, écart-type relatif, coefficient de variation.

#### II.4.2 Ecart-type

L'écart-type symbolisé par  $s$  ( $N \leq 20$ ) ou  $\sigma$  ( $N > 20$ ) désigne l'erreur aléatoire ou fortuite d'une série de mesures d'un même échantillon. En chimie analytique cette grandeur nous permet de nous faire une idée sur la méthode de travail utilisée. C'est donc  $s$  ( $N \leq 20$ ) ou  $\sigma$  ( $N > 20$ ) qui fait foi quand il faudra se prononcer sur la qualité de la méthode d'analyse utilisée. Plus  $s$  sera petit, meilleur sera la méthode choisie. Il va de soi, qu'en répétant les mesures sur le même échantillon, il faut s'attendre à des fluctuations quant à la répétitivité de  $s$ . Cela est principalement dû au jeu du hasard parfaitement assimilable à l'analyse quantitative. Ce n'est qu'en exerçant un nombre infini de mesures sur un échantillon que l'on pourra se rapprocher de la vraie valeur de  $s$  qui devient alors  $\sigma$  car  $N \gg 20$ . Comme dans la pratique, on ne peut pas se permettre d'analyser chaque échantillon une vingtaine de fois, nous devons nous contenter d'un nombre restreint de manipulations. Toute notre information est basée sur un échantillonnage relativement petit. La question qui se pose alors est la suivante : est-il possible de se prononcer fermement sur la qualité de la mesure et sur ses erreurs?

La réponse est affirmative et l'on peut se prononcer avec une certaine aisance sur le résultat même avec un nombre limité de mesures et d'échantillons. Dès lors, on doit introduire la notion de niveau de confiance et de degré d'incertitude. La valeur déterminée sera alors comprise entre une valeur d'incertitude inférieure et supérieure. Ce degré d'incertitude est de lui-même fonction du nombre de degrés de liberté du système. Pour ce faire, on utilise les tables de distribution des lois de Student, test-F..., selon le cas. Ce dernier point sera étudié en détail au cours de votre cursus universitaire. Les points mesurés sont groupés autour d'une valeur moyenne définie par la loi de gauss. Pour calculer l'écart-type  $s$  ( $N \leq 20$ ) ou  $\sigma$  ( $N > 20$ ) on utilise :

$$s_x = \sqrt{\frac{\sum_{i=1}^{N-1} (x_i - \bar{x})^2}{N-1}} \quad \text{et} \quad \sigma_x = \sqrt{\frac{\sum_{i=1}^{N-1} (x_i - \mu_x)^2}{N}}$$

Avec  $N-1$  : le nombre de degré de liberté du système pour ( $N \leq 20$ )

### II.4.3 Ecart-type relatif et coefficient de variation

On peut aussi, à partir de l'écart-type définir un écart-type relatif ("relative standard deviation" en anglais)  $s_r$  ( $N \leq 20$ ) ou  $\sigma_r$  ( $N > 20$ ) et le coefficient de variation  $CV$ :

$$s_{rx} = \frac{s_x}{\bar{x}} \text{ ou } \sigma_{rx} = \frac{\sigma_x}{\mu_x} \text{ et } CV_x = \frac{s_x}{\bar{x}} \times 100 \text{ ou } CV_x = \frac{\sigma_x}{\mu_x} \times 100$$

Attention, ces grandeurs ainsi définies sont adimensionnelles, ce qui n'est pas le cas de l'écart-type.

## II.5 Qualité d'analyse

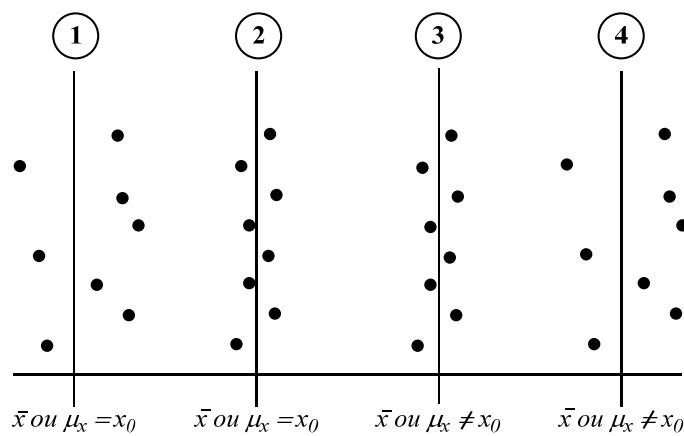
### II.5.1 Expression du résultat analytique

Dans tous les cas de figure, un résultat analytique doit être exprimé par une moyenne et un écart-type. Les erreurs systématiques, qui n'agissent que sur la moyenne des résultats, ne doit pas apparaître. On doit présenter la moyenne corrigée. La moyenne n'est corrigée que dans le cas où on a réussi à déterminer précisément les erreurs systématiques par recalibration de l'appareil de mesure par exemple. On aura donc pour un échantillon :

$$\bar{x} \pm s_x \text{ ou } \mu_x \pm \sigma_x$$

### II.5.2 Comparaison entre les différentes sources d'erreur

Voici un schéma synoptique qui illustre ces concepts :



Cas n°1 : juste et peu précis ; Cas n°2 : juste et précis ; Cas n°3 : pas juste et précis ; Cas n°4 : pas juste et peu précis

Dans ces cas de figure, le meilleur résultat est **2** et le moins bon est **4**.

#### II.5.3 Sensibilité de la méthode analytique

La précision d'une mesure peut être limitée par la sensibilité de la méthode analytique choisie. Elle peut être exprimée en écart-type, écart-type relatif ou coefficient de variation (le pouvoir de résolution). La sensibilité d'une méthode analytique est définie à la fois par la sensibilité de l'appareillage de mesure et par sa réponse à un analyte donné. Par exemple, en gravimétrie, on obtiendra une pesée moyenne de  $500 \text{ mg} \pm 2 \text{ mg}$  avec une balance qui n'est sensible qu'à 1 mg, la masse ne pourra donc être connue qu'à 1 mg près au maximum. En sus de la sensibilité de l'appareillage vient se greffer la limite de détection (LOD) et la limite de quantification (LOQ) de l'analyte en question. Ces deux grandeurs sont déterminées à l'aide de la courbe d'étalonnage de l'analyte étudié obtenue sur l'appareil analytique en question. Dès lors, la sensibilité pourra être définie pour un couple appareil analytique-analyte donné. La détermination de ces deux grandeurs sera abordée dans le chapitre relatif à l'étalonnage des appareils analytiques.

#### II.5.4 Exemple n°1

Quatre étudiants ont fait le titrage de 10,00 mL d'une solution de NaOH ( $0,100 \text{ mol} \cdot \text{L}^{-1}$ ) par HCl ( $0,100 \text{ mol} \cdot \text{L}^{-1}$ ). Le calcul nous annonce une valeur de 10,00 mL de NaOH. Les étudiants ont obtenu les valeurs suivantes :

Étudiant	$V_{\text{HCl}} (1)$	$V_{\text{HCl}} (2)$	$V_{\text{HCl}} (3)$	$V_{\text{HCl}} (4)$	$V_{\text{HCl}} (5)$
1	10,08	10,11	10,09	10,10	10,12
2	9,88	10,14	10,02	9,80	10,21
3	10,19	9,79	9,69	10,05	9,78
4	10,04	9,98	10,02	9,97	10,04

Etudiant	$x_i$ (mL)	$\bar{x}$ (mL)	$(x_i - \bar{x})$ (mL)	$(x_i - \bar{x})^2$ (mL <sup>2</sup> )
1	10,08		-0,02	0,0004
	10,11		0,01	0,0001
	10,09	10,10	-0,01	0,0001
	10,10		0,00	0,0000
	10,12		0,02	0,0004
<i>Sommes</i>	50,50		0	0,0010
2	9,88		-0,13	0,0169
	10,14		0,13	0,0169
	10,02	10,01	0,01	0,0001
	9,80		-0,21	0,0441
	10,21		0,20	0,0400
<i>Sommes</i>	50,05		0	0,1180
3	10,19		0,29	0,0841
	9,79		-0,11	0,0121
	9,69	9,90	-0,21	0,0441
	10,05		0,15	0,0225
	9,78		-0,12	0,0144
<i>Sommes</i>	49,50		0	0,1772
4	10,04		0,03	0,0009
	9,98		-0,03	0,0009
	10,02	10,01	0,01	0,0001
	9,97		-0,04	0,0016
	10,04		0,03	0,0009
<i>Sommes</i>			0	0,0044

Pour l'étudiant 1 le détail des calculs est le suivant :

$$\text{Moyenne : } \overline{V_{HCl}} = \frac{\sum_{i=1}^{i=5} V_{HCl}}{5} = \frac{5,50}{5} = 10,10 \text{ mL}$$

$$\text{Écart type : } S_{V_{HCl}} = \sqrt{\frac{\sum_{i=1}^{i=5} (V_{i(HCl)} - \overline{V_{HCl}})^2}{5-1}} = \sqrt{\frac{0,0010}{4}} = 0,0158 \text{ mL}$$

Pour cet étudiant, il conviendra d'écrire le résultat comme étant :  $\overline{V_{HCl}} = 10,10 \pm 0,016 \text{ mL}$ .

Attention, ici nous n'avons pris en compte que la variation du volume lu sur la burette, il s'agit de l'erreur de mesure. Pour être totalement rigoureux, l'écart-type global doit prendre en compte aussi toutes les autres erreurs issues de la dilution de l'échantillon (pipette, ballon jaugé, précision de la burette) ainsi que l'erreur sur la solution de titrant. Dans ce cas, on effectue une étude de propagation d'erreur que l'on étudiera en détail plus loin. En effectuant ce même genre de calcul pour chaque étudiant, il vient :

Étudiant	$\overline{V_{HCl}}$ (mL)	$s_{V_{HCl}}$ (mL)	$E_r \overline{V_{HCl}} (V_{0(HCl)})$ (%)	Conclusion
1	10,10	0,016	1	Peu juste et très précis
2	10,01	0,172	0,1	Très juste et peu précis
3	9,90	0,210	1	Peu juste et très peu précis
4	10,01	0,033	0,1	Très juste et précis

### III Propagation des erreurs à travers une chaîne de mesures

#### III.1 Introduction

En chimie analytique, il faut souvent estimer l'erreur systématique d'un résultat qui a été calculé à partir de plusieurs données expérimentales, ayant chacune une erreur systématique connue. Dans bien des cas, si l'appareil de mesure a bien été calibré et qu'il n'y a pas eu d'erreurs systématiques dans la préparation du calibrant et/ou des échantillons à analyser, cette opération n'a pas lieu d'être. Dans le cas contraire, l'erreur systématique du résultat final dépend des opérations mathématiques réalisées entre les différentes données expérimentales. Il est alors nécessaire d'effectuer un calcul combinant toutes les erreurs systématiques identifiées afin de corriger le résultat analytique final. De même, il faudra combiner les erreurs aléatoires attenantes à la chaîne de mesure qui elles ne peuvent être corrigées et qui seront reportées avec le résultat de mesure proposé.

## III.2 Propagation des erreurs à travers une chaîne de mesures

### III.2.1 Propagation des erreurs systématiques

L'erreur systématique sur chaque opération d'une chaîne de mesure doit être déterminée. Toutes ces erreurs doivent être combinées, en fonction de la chaîne de mesure mise en œuvre, afin que le résultat lié à l'enchaînement des diverses opérations soit corrigé. Si la correction sur le résultat final n'est pas apportée, alors on doit considérer le biais global de la méthode de détermination. Une méthode valide doit proposer un biais minimal et dans le cas absolu aucun biais. Ci-dessous un tableau qui retrace les différentes combinaisons d'erreurs systématiques liées à une chaîne de mesures :

Opération	Exemple	Ecart-type du résultat final
Addition Soustraction	$\bar{y} = \bar{a} + \bar{b} - \bar{c}$	$\Delta y = \Delta a + \Delta b + \Delta c$
Combinaison linéaire	$\bar{y} = k_a \bar{a} + k_b \bar{b} - k_c \bar{c}$	$\Delta y = k_a \Delta a + k_b \Delta b + k_c \Delta c$
Multiplication Division	$\bar{y} = \frac{(\bar{a} \times \bar{b})}{\bar{c}}$	$\frac{\Delta y}{y} = \left( \frac{\Delta a}{\bar{a}} \right) + \left( \frac{\Delta b}{\bar{b}} \right) + \left( \frac{\Delta c}{\bar{c}} \right)$
Fonction	$\bar{y} = f(\bar{x})$ $\bar{G} = f(\bar{x}, \bar{y}, \bar{z})$	$\Delta y = \Delta x \left  \frac{dy}{dx} \right $ $\Delta G = \Delta x \left  \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial x} \right  + \Delta y \left  \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial y} \right  + \Delta z \left  \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial z} \right $

Dans ce tableau,  $k_a$ ,  $k_b$  et  $k_c$  sont des constantes.  $\bar{a}, \bar{b}, \bar{c}, \bar{x}, \bar{y}$  et  $\bar{z}$  représentent des moyennes. Il en résultera une erreur systématique globale, engendrant un biais rapport à la valeur vraie :  $biais = \Delta \bar{y}(\bar{y}_0) = \bar{y} - \bar{y}_0$ .

### III.2.2 Propagation des erreurs aléatoires

L'erreur aléatoire sur chaque opération d'une chaîne de mesure doit être déterminée. Toutes ces erreurs doivent être combinées, en fonction de la chaîne de mesure mise en œuvre, afin

que la précision globale du résultat lié à l'enchaînement des diverses opérations soit reportée aux côtés de la moyenne de la détermination. Ci-dessous un tableau qui retrace les différentes combinaisons d'erreurs aléatoires liées à une chaîne de mesures :

Opération	Exemple	Ecart-type du résultat final
Addition Soustraction	$\bar{y} = \bar{a} + \bar{b} - \bar{c}$	$s_y = \sqrt{s_a^2 + s_b^2 + s_c^2}$
Combinaison linéaire	$\bar{y} = k_a \bar{a} + k_b \bar{b} - k_c \bar{c}$	$s_y = \sqrt{(k_a s_a)^2 + (k_b s_b)^2 + (k_c s_c)^2}$
Multiplication Division	$\bar{y} = \frac{(\bar{a} \times \bar{b})}{\bar{c}}$	$\frac{s_y}{\bar{y}} = \sqrt{\left(\frac{s_a}{\bar{a}}\right)^2 + \left(\frac{s_b}{\bar{b}}\right)^2 + \left(\frac{s_c}{\bar{c}}\right)^2}$
Fonction	$\bar{y} = f(\bar{x})$ $\bar{G} = f(\bar{x}, \bar{y}, \bar{z})$	$s_y = s_x \left  \frac{dy}{dx} \right $ $s_G = \sqrt{\left( \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial \bar{x}} \right)^2 \times s_x^2 + \left( \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial \bar{y}} \right)^2 \times s_y^2 + \left( \frac{\partial f(\bar{x}, \bar{y}, \bar{z})}{\partial \bar{z}} \right)^2 \times s_z^2}$

Dans ce tableau,  $k_a$ ,  $k_b$  et  $k_c$  sont des constantes.  $\bar{a}, \bar{b}, \bar{c}, \bar{x}, \bar{y}$  et  $\bar{z}$  représentent des moyennes. Il en résultera une erreur aléatoire globale. Si cette dernière est plus grande que l'erreur systématique globale alors le biais de la méthode d'analyse ne doit pas être considéré.

### III.3 Exemples

#### III.3.1 Exemple n°2

Un échantillon de 3,4842 g d'un mélange solide qui contient de l'acide benzoïque ( $C_6H_5COOH$ ,  $M = 122,123 \text{ g} \cdot \text{mol}^{-1}$ ) a été dissous et titré par une base en présence de phénolphtaléine. Au point d'équivalence on a jouté 41,36 mL de  $\text{NaOH}$  ( $0,2328 \text{ mol} \cdot \text{L}^{-1}$ ). Quel est la fraction massique de l'acide benzoïque dans le mélange solide ?

À l'équivalence, l'acide benzoïque étant un mono-acide on aura :

$$\overline{n_{C_6H_5COOH}} = \overline{n_{NaOH}} = \overline{C_{NaOH}} \times \overline{V_{NaOH}}$$

Or la masse de C<sub>6</sub>H<sub>5</sub>COOH est donnée par :

$$\overline{m_{C_6H_5COOH}} = \overline{n_{C_6H_5COOH}} \times \overline{M_{C_6H_5COOH}}$$

On aura donc que le pourcentage massique sera donné par :

$$\frac{\overline{m_{C_6H_5COOH}}}{\overline{m_{solide}}} \times 100 = \frac{\overline{C_{NaOH}} \times \overline{V_{NaOH}} \times \overline{M_{C_6H_5COOH}}}{\overline{m_{solide}}} \times 100 = 33,749 \%$$

On a mesuré les incertitudes suivantes :

Incertitude sur la lecture de la burette :  $\pm 0,02 \text{ mL}$

Incertitude sur la masse :  $\pm 0,0001 \text{ g}$

Incertitude sur la concentration de NaOH :  $\pm 0,0001 \text{ mol} \cdot \text{L}^{-1}$

Incertitude sur la masse molaire : négligeable

Pour déterminer l'incertitude sur le pourcentage massique en acide benzoïque, il faut penser que l'on a fait deux lectures de burette et donc deux erreurs, une à la mise à zéro et une à la lecture du volume d'équivalence :

$$S_{V_{NaOH}} = \sqrt{(0,02)^2 + (0,02)^2} = 0,028 \text{ mL}$$

$$\frac{s_{\%}}{\%} = \sqrt{\left(\frac{S_{C_{NaOH}}}{C_{NaOH}}\right)^2 + \left(\frac{S_{V_{NaOH}}}{V_{NaOH}}\right)^2 + \left(\frac{S_{m_{solide}}}{m_{solide}}\right)^2} = \sqrt{\left(\frac{0,0001}{0,2328}\right)^2 + \left(\frac{0,028}{41,36}\right)^2 + \left(\frac{0,0001}{3,4842}\right)^2} = 8,02 \times 10^{-4}$$

On aura alors :

$$s_{\%} = 8,02 \times 10^{-4} \times \sqrt{\%} = 0,0270$$

On aura donc un pourcentage d'acide benzoïque dans le mélange solide de  $33,75\% \pm 0,03\%$

Dans les exemples que nous avons vus précédemment, l'incertitude du résultat ne dépendait que de l'incertitude liée à la préparation de l'échantillon à analyser et à la précision de la burette. C'est typiquement le cas d'un dosage par indicateur coloré où l'incertitude liée à la mesure elle-même n'est pas quantifiable. En effet, le virage d'un indicateur n'est pas une grandeur quantitative, mais une grandeur qualitative qui dépendent de l'acuité visuelle de l'expérimentateur. Nous allons dès à présent nous intéresser au cas où l'on peut quantifier

l'incertitude liée à la mesure d'une grandeur. Ceci n'est possible que lorsque l'on utilise un appareil de mesure (pH-mètre, voltmètre...), ce qui est le cas le plus répandu en chimie analytique. Dans cette optique, on doit définir une incertitude totale sur la mesure d'une grandeur comme le pH par exemple. L'incertitude totale doit prendre en compte l'incertitude sur la mesure de l'échantillon analysé et l'incertitude sur la préparation de cet échantillon. D'un point de vue écart-type on aura :

$$s_{pH} = \sqrt{(s_{pH_{mes}})^2 + (s_{pH_{prép}})^2}$$

Dans le cas où l'incertitude sur la mesure est plus grande (au moins 100 fois) que l'incertitude sur la préparation, alors cette dernière est négligeable et donc  $s_{pH} \approx s_{pH_{mes}}$ .

### III.3.2 Exemple n°3

Pour mesurer le pH d'une solution, on doit étalonner le pH-mètre. Après étalonnage, pour vérifier la calibration, on prépare 1000 mL d'une solution test à pH 2. Nous avons à disposition de l'acide chlorhydrique concentré à  $1 \text{ mol} \cdot \text{L}^{-1} \pm 0,0001 \text{ mol} \cdot \text{L}^{-1}$ . A partir de cette solution d'acide concentré, il faut préparer une solution à  $1 \cdot 10^{-2} \text{ mol} \cdot \text{L}^{-1}$  en diluant par 100, c'est-à-dire introduire 10 mL d'acide chlorhydrique dans la fiole jaugée et compléter jusqu'au trait de jauge. La mesure du pH de notre solution test donne les valeurs suivantes : 1,90 ; 2,00 ; 2,20 ; 1,80 ; 2,10.

Calculer l'incertitude sur la mesure du pH.

On a :

Solution d'acide chlorhydrique :  $\pm 0,0001 \text{ mol} \cdot \text{L}^{-1}$

Pipette de 10 mL :  $\pm 0,02 \text{ mL}$

Fiole jaugée de 1000 mL :  $\pm 0,40 \text{ mL}$

Sachant que l'incertitude sur le pH dépend de la préparation de la solution test et de la mesure de cette dernière, il faut calculer les deux incertitudes. Pour ce faire, il faut d'abord calculer le pH moyen mesuré :

$$\text{pH moyen : } \overline{\text{pH}} = \frac{1,90 + 2,00 + 2,20 + 1,80 + 2,10}{5} = 2,00$$

Incertitude sur la mesure du pH :

$$s_{pH_{mes}} = \sqrt{\frac{(1,90-2,00)^2 + (2,00-2,00)^2 + (2,20-2,00)^2 + (1,80-2,00)^2 + (2,10-2,00)^2}{4}} = 0,16$$

Incertitude sur la préparation de la solution test :

$$pH = -\log\left(\frac{[H_3O^+]}{C^0}\right) \text{ et } s_{pH_{prép}} = 0,434 \frac{s_{C_{H_3O^+}}}{C_{H_3O^+}}$$

$$\overline{C_{H_3O^+}} = \frac{\overline{C_{HCl}} \times \overline{V_{HCl}}}{\overline{V_{solution}}}$$

D'où :

$$\frac{s_{C_{H_3O^+}}}{\overline{C_{H_3O^+}}} = \sqrt{\left(\frac{s_{C_{HCl}}}{\overline{C_{HCl}}}\right)^2 + \left(\frac{S_{V_{HCl}}}{\overline{V_{HCl}}}\right)^2 + \left(\frac{S_{V_{solution}}}{\overline{V_{solution}}}\right)^2} = \sqrt{\left(\frac{0,0001}{1}\right)^2 + \left(\frac{0,02}{10}\right)^2 + \left(\frac{0,40}{1000}\right)^2} = 0,002$$

On a alors :

$$s_{pH_{prép}} = 0,434 \frac{s_{C_{H_3O^+}}}{C_{H_3O^+}} = 0,434 \times 0,002 = 0,00087$$

L'incertitude totale sur le pH de la solution test est :

$$s_{pH} = \sqrt{(s_{pH_{mes}})^2 + (s_{pH_{prép}})^2} = \sqrt{(0,16)^2 + (0,00087)^2} = 0,1600$$

Attendu que le pH d'une solution se donne à deux chiffres après la virgule alors quel que soit la formule utilisée on aura :

$$pH = 2,00 \pm 0,16$$

On remarque à travers cet exemple que l'incertitude sur la mesure est souvent beaucoup grande que l'incertitude sur la préparation de l'échantillon.

## **IV Chiffres significatifs**

### **IV.1 Introduction**

Une manière simple d'indiquer le degré d'incertitude associé à une mesure expérimentale est d'arrondir le résultat afin qu'il ne contienne que les chiffres significatifs.

Par exemple, la lecture d'une burette est estimable à  $\pm 0,02$  mL. Ainsi, le volume devra être donné avec 4 chiffres significatifs comme 35,56 mL par exemple.

### **IV.2 Chiffres significatifs de données brutes**

Par exemple, considérons l'ensemble des résultats : 61,60 ; 61,46 ; 61,55 ; 61,61. En calculant la moyenne et l'écart-type, on trouve :  $61,555 \pm 0,069$ . Le résultat arrondi sera :  $61,56 \pm 0,07$ . Lorsqu'un nombre finit par 5, il faut toujours arrondir au nombre pair le plus proche.

### **IV.3 Chiffres significatifs de données calculées à partir de données brutes**

#### **IV.3.1 Sommes et différences**

Pour l'addition ou la soustraction, on peut trouver le nombre de chiffres significatifs facilement :

$$3,4 + 0,0020 + 7,31 = 10,73 = 10,7$$

Ce résultat possède alors 3 chiffres significatifs.

#### **IV.3.2 Produits et quotients**

Pour la multiplication et la division, on a souvent tendance à arrondir le résultat avec le même nombre de chiffres significatifs que le terme qui en possède le moins. Regardons à présent ces deux exemples où les résultats sont exprimés selon la réponse que donne la calculatrice :

$$\frac{24 \times 4,52}{100,0} = 1,0848 \quad \text{et} \quad \frac{24 \times 4,02}{100,0} = 0,9648$$

Selon la règle, les résultats devraient être :

$$1,1 \text{ et } 0,96$$

En effet, le premier résultat possède 2 chiffres significatifs comme 24 et le deuxième résultat possède lui aussi 2 chiffres significatifs car les zéro qui ne servent qu'à localiser la virgule ne sont pas significatifs. Maintenant examinons l'incertitude sur les chiffres de l'opération. L'incertitude relative sur ces nombres est de 1 pour 24 ; 0,01 pour 4,52 et 0,1 pour 100. L'incertitude relative sur le résultat est donné par :

$$s = \sqrt{\left(\frac{1}{24}\right)^2 + \left(\frac{0,01}{4,52}\right)^2 + \left(\frac{0,1}{100}\right)^2} = 0,04$$

L'incertitude relative étant  $\pm 0,04$ , le résultat devra être 1,08 et non 1,1. Pour le deuxième résultat on a aussi une incertitude relative de  $\pm 0,04$ . Le résultat devra être 0,96.

#### IV.3.3 Logarithmes et exponentielles

Pour le logarithme d'un nombre, on garde autant de chiffres après la virgule qu'il n'y a de chiffres significatifs dans le nombre de départ :

$$\log(6,000 \times 10^{-5}) = -4,2218$$

Pour l'exponentielle ou puissance d'un nombre, on conserve autant de chiffres qu'il y a de chiffres après la virgule dans le nombre de départ :

$$10^{12,5} = 3 \times 10^{12}$$