

# L'apport des méthodes statistiques dans la maîtrise de la qualité des analyses

A. Jardy\* et J. Vial

Laboratoire « Environnement et Chimie Analytique », associé au CNRS, École Supérieure de Physique et Chimie Industrielles de la Ville de Paris, 10 rue Vauquelin, 75005 Paris, France

It seems an evidence nowadays that a statistical approach is quite necessary in modern analytical chemistry both for data processing and design of experiments. In spite of technical progress, which has enabled full automation of the whole analysis sequences, differences are often observed between repeatability and reproducibility which are accurately defined. Any rigorous reproducibility study requires the implementation of a collaborative study, and the general technique used for data evaluation is the analyses of variance (ANOVA). In its one-way version, the influence of the laboratory factor can be studied globally and the variance due to the between-laboratories source of variation estimated if it is found significantly different from zero. In order to minimize any dispersive effect, it is necessary to identify among the sources of error involved, which ones have a significant influence: ANOVA is again the best method of investigation. As in quantitative determinations the true value is unknown, the challenge is to give the best estimation through a confidence interval. In data exploitation, the problem arise of the discrimination between random errors and systematic errors and their propagation. The greatest difficulty in accuracy studies is encountered when no standards nor certified reference materials are available; in this case, only indirect methods can be carried out.

Il serait inconcevable, dans un dossier « Métrologie en Chimie » de ne pas examiner l'apport des méthodes statistiques aussi bien pour l'exploitation des résultats que pour l'organisation des expériences. Il existe en effet, lors de toute détermination expérimentale, des sources de variations multiples qui tendent à introduire une dispersion des résultats autour de la valeur vraie. Celle-ci étant par essence même inconnue, l'enjeu est d'en donner la meilleure estimation possible. Cette approche implique donc de disposer d'une méthodologie apte à identifier les causes d'erreur, à en déterminer les effets pour enfin les minimiser par une optimisation raisonnée des conditions opératoires.

## Introduction

Les exigences croissantes du contrôle qualité, que l'on rencontre dans maints secteurs industriels de pointe, tels l'industrie pharmaceutique, l'industrie nucléaire ou la chimie de l'environnement, requièrent une maîtrise accrue de la qualité des analyses produites. Or il est bien connu qu'un même problème d'analyse peut conduire à des résultats différents [1-2] s'il est traité selon un même protocole sur des sites différents, *a fortiori* si chaque laboratoire impliqué a le libre choix de la méthode mise en œuvre. Prenons l'exemple de la détermination de l'atrazine, herbicide de la famille des triazines, dans un échantillon d'eau, ayant fait l'objet d'une étude interlaboratoires [3]. À l'examen des résultats (Fig. 1)

où chaque laboratoire a effectué trois analyses de cette solution, utilisant la méthode de son choix, il apparaît :

- une dispersion non homogène des résultats de déterminations répétées au sein d'une même unité ;
- une dispersion de la teneur en atrazine déterminée expérimentalement d'un laboratoire à l'autre ; ainsi l'écart relatif entre la moyenne des résultats dans le laboratoire n° 14 et la moyenne de l'ensemble des laboratoires dépasse 30 %.

Cet exemple illustre bien que dans l'évaluation de la qualité des analyses, il faut faire intervenir :

- la répétabilité, caractérisant la dispersion intralaboratoire au cours d'une même série d'essais, c'est-à-dire entre analyses répétées par le même opérateur, utilisant le même appareillage sur un court intervalle de temps ;
- la reproductibilité, caractérisant la dispersion des résultats observés dans des laboratoires distincts, donc par des opérateurs distincts utilisant des appareillages différents mais appliquant le même protocole, ce qui revient à prendre en compte les fluctuations interlaboratoires ;
- la justesse, qui traduit l'absence d'écart significatif entre le résultat proposé – les statisticiens disent la valeur estimée – et la valeur vraie de la concentration de l'analyte, en d'autres termes, l'absence d'erreur systématique – ou biais d'analyse – (Fig. 2).

De fait, l'étude de la justesse d'une méthode, ou d'une chaîne métrologique, c'est-à-dire la recherche d'erreurs systématiques éventuelles passe par la mise en œuvre d'étalons, ou à tout le moins de matériaux de référence certifiés (CRM) [4-6]. Cela conduit dès lors à la recherche des réponses les

plus objectives possibles à des questions du type : « l'écart entre la valeur estimée et la valeur attendue peut-il simplement s'expliquer en termes de dispersion des résultats, lié au nombre limité de déterminations - en termes d'aléas d'échantillonnage - ou, ceux-ci n'expliquant pas tout, faut-il conclure à l'existence d'un biais ? » ou « le degré de compatibilité entre des résultats obtenus avec des méthodes ou des chaînes différentes est-il satisfaisant ? »

Il en découle que dans toute perspective de maîtrise de la qualité et afin de donner un maximum d'efficacité aux tests de justesse, il faut viser dans un premier temps à réduire la dispersion des analyses, ce qui implique d'identifier les sources de variations, d'en quantifier les effets pour enfin les minimiser par un choix raisonné des conditions expérimentales [7].

### Répétabilité

La mise en œuvre de répétitions, pour chaque détermination de la teneur en analyte dans un échantillon donné présente deux avantages :

- premièrement des résultats entachés d'erreurs grossières peuvent être détectés et éliminés, si un test statistique permet de les considérer comme des valeurs aberrantes [8] ;
- deuxièmement l'incidence des erreurs accidentelles, qui traduisent l'effet de tous les facteurs non contrôlés, multiples dans toute méthode analytique, se trouve minimisé : si l'écart type de la distribution des résultats d'une détermination isolée est  $\sigma_r$ , celui de la moyenne de  $n$

répétitions est divisé par  $\sqrt{n}$ , soit  $\frac{\sigma_r}{\sqrt{n}}$ .

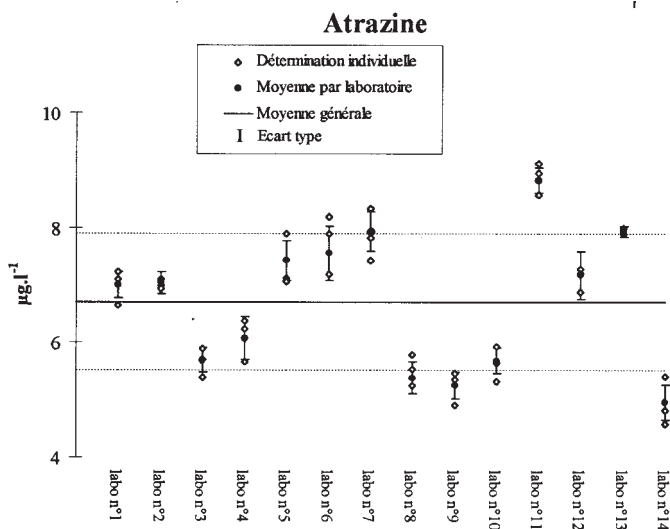


Figure 1. Teneur en atrazine observée lors du circuit interlaboratoires (d'après [3]).

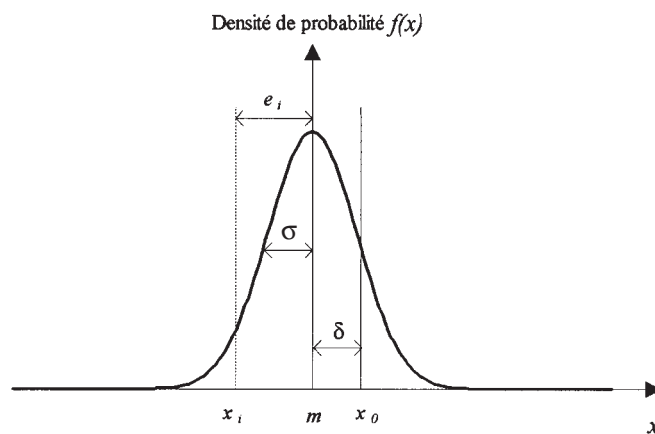


Figure 2. Caractéristiques de distribution des résultats de mesure.  $x_0$  : valeur vraie,  $m$  : valeur la plus probable (espérance mathématique),  $x_i$  : valeur observée à la  $i^{\text{ème}}$  mesure,  $e_i$  : erreur accidentelle (aléatoire),  $\delta$  : erreur systématique (biais),  $\sigma$  : écart type.

Dans l'hypothèse d'une distribution suivant une loi normale, l'intervalle ayant une même probabilité  $(1 - \alpha)$  de contenir la valeur observée, dit intervalle de confiance de seuil  $\alpha$ , est beaucoup plus étroit pour la moyenne de  $n$  répétitions que pour une détermination unique (Fig. 3).

L'ensemble des résultats possibles pour une même détermination expérimentale constitue par essence même une population infinie,  $\sigma_r$  ne peut être atteint avec certitude, mais seulement estimé. La préoccupation pratique de l'analyste est alors d'utiliser au mieux l'information disponible, pour en donner la meilleure estimation possible  $\hat{\sigma}_r$  à partir du plus petit nombre de mesures.

Deux cas sont alors à considérer :

- $\sigma_r$  doit être estimé à partir d'une seule série de  $n$  répétitions, pour une seule concentration de l'analyte. On utilise alors classiquement la relation (1) donnant une estimation dite à  $(n - 1)$  degrés de liberté. Ce nombre de degrés de liberté constitue en quelque sorte une quantification de l'information collectée.

$$\hat{\sigma}_r = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n - 1}} \quad (1)$$

- $\sigma_r$  peut être estimé à partir des  $N$  résultats obtenus pour  $k$  échantillons de teneur en analyte analogues, un test statistique ayant permis de conclure à l'homogénéité de la variance (test de Bartlett, de Cochran, de Hartley... décrits dans les ouvrages statistiques [8-9]). Cette

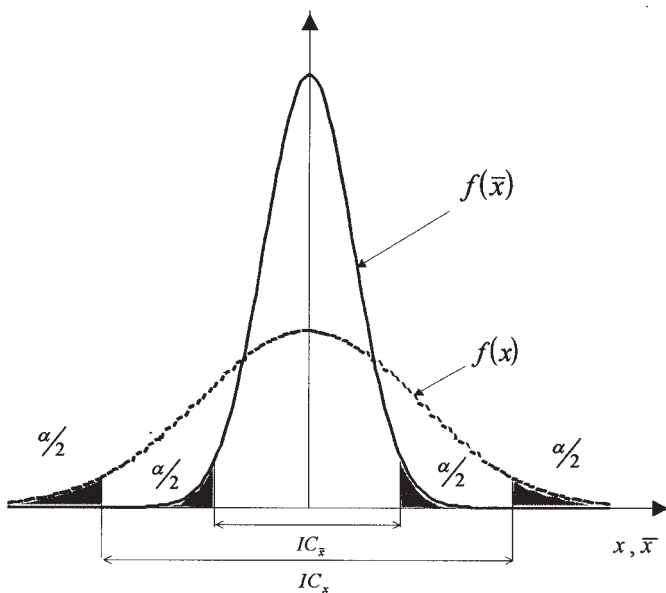


Figure 3. Différences entre les distributions d'une variable  $x$  et de sa moyenne  $\bar{x}$ , ici  $n = 6$ .

variance commune aux  $k$  populations (chaque solution de l'analyte générant une population de résultats possibles par application de la méthode considérée) dite variance intraclasse est estimée par (2) :

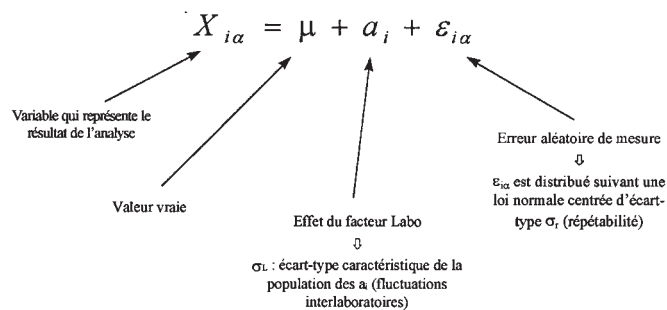
$$\hat{\sigma}_r^2 = \frac{\sum_i (x_{1i} - \bar{x}_1)^2 + \sum_i (x_{2i} - \bar{x}_2)^2 + \dots + \sum_i (x_{ki} - \bar{x}_k)^2}{N - k} \quad (2)$$

c'est-à-dire par cumul des sommes des carrés des écarts à la moyenne des résultats pour chaque solution. L'estimation  $\hat{\sigma}_r$  de l'écart type correspondant comporte alors  $(N - k)$  degrés de liberté, indice d'une plus grande fiabilité.

### Reproductibilité

L'étude de la reproductibilité d'une méthode d'analyse ne peut se faire que par la mise en œuvre d'un circuit interlaboratoires : le même échantillon contenant l'analyte est distribué à  $p$  laboratoires (un laboratoire est défini comme un couple opérateur-appareillage), à qui il est demandé d'effectuer  $n$  répétitions de la même analyse.

La méthode statistique utilisée pour l'exploitation des résultats est l'analyse de la variance (ANOVA), ici à un facteur contrôle, l'identité du laboratoire. Ce facteur est dit comporter  $p$  modalités. Le modèle mathématique à la base du dépouillement est :



Pour déterminer, à l'aide d'un test statistique, si l'influence du facteur contrôlé - l'identité du laboratoire - est mise en évidence, on compare la dispersion entre les moyennes observées pour chaque modalité de ce facteur autour de la moyenne générale à celle observée entre les répétitions pour chaque modalité. Il est clair que si le facteur contrôlé a une influence sur la réponse mesurée, ce qui correspond dans le modèle à considérer que tous les  $a_i$  ne sont pas identiquement nuls, la dispersion interlaboratoires ne peut être que supérieure à celle due aux seuls aléas d'échantillonnage (Fig. 4). On est alors amené à introduire un terme dispersif supplémentaire sous forme d'un écart-type de fluctuation interlaboratoires  $\sigma_L$  dont l'estimation repose sur l'hypothèse que le panel de laboratoires ayant

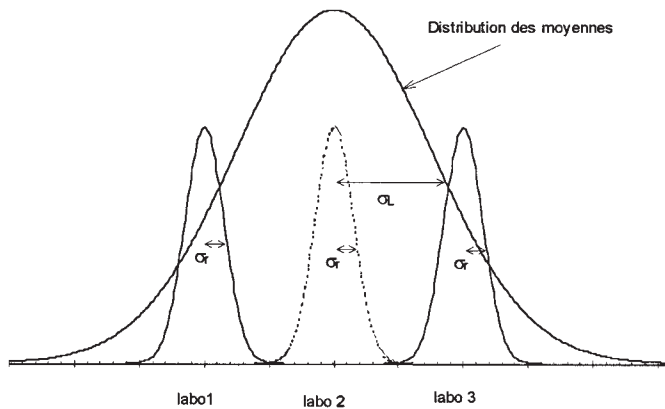


Figure 4. Dispersions intralaboratoire et interlaboratoires.

participé au circuit est représentatif de la population des laboratoires susceptibles de mettre en œuvre cette méthode.

Le tableau théorique d'analyse de la variance est représenté tableau I.

Il apparaît que pour tester l'absence d'influence du facteur contrôlé, laquelle se traduirait par  $\sigma_L = 0$ , il suffit d'observer que sous réserve que cette hypothèse soit vérifiée,  $q_L$  et  $q_r$  sont deux estimations indépendantes de  $\sigma_r^2$ . La réponse à la question posée est donc fournie par un test de Snedecor de comparaison de deux variances estimées, avec respectivement  $(p - 1)$  et  $(N - p)$  degrés de liberté. Si la valeur observée appartient au groupe de celles jugées d'obtention très peu probable (probabilité  $\leq \alpha$ ) dans le cadre de l'hypothèse faite, on préfère remettre en cause celle-ci et partant conclure à son rejet, avec un risque d'erreur égal à  $\alpha$ . En

pratique cela revient à comparer la valeur observée pour le rapport  $\frac{q_L}{q_r}$  à la valeur limite qu'une variable de Snedecor à

$(p - 1)$  et  $(N - p)$  degrés de liberté a  $(1 - \alpha)$  chances sur 100 de ne pas dépasser. Dans le cas où l'hypothèse  $\sigma_L = 0$  peut être rejetée, et dans ce cas seulement,  $\sigma_L$  peut être

estimé par  $\hat{\sigma}_L = \sqrt{\frac{q_L - q_r}{n}}$ . Il en résulte que si l'on conclut

$\sigma_L \neq 0$ , la reproductibilité diffère de la répétabilité et est caractérisée par un écart-type  $\sigma_R$  estimé par  $\hat{\sigma}_R = \sqrt{\hat{\sigma}_L^2 + \frac{\hat{\sigma}_r^2}{n}}$ .

On notera que dans ce cas, il existe une limitation pratique à une augmentation du nombre de répétitions, dès lors que  $\frac{\hat{\sigma}_r^2}{n}$  devient nettement inférieur à  $\hat{\sigma}_L^2$ . La qualité de tout

résultat analytique implique donc une caractérisation complète de la méthode, dès sa phase développement et validation. En outre, l'amélioration de la qualité des analyses passe par une diminution de  $\sigma_r$  aussi bien que de  $\sigma_L$ , ce qui implique l'identification des sources d'erreur.

### Recherche des sources d'erreur

Il existe de multiples sources d'erreur lors de la mise en œuvre de toute méthode d'analyse, depuis l'échantillonnage jusqu'à l'analyse elle-même, en passant par la préparation et le traitement des échantillons [10-11]. À titre d'exemple, des causes d'erreur possibles lors de l'étape d'analyse par une méthode chromatographique, on peut citer :

- l'identification fondée sur la seule mesure des temps de rétention ;

Tableau I. Tableau théorique d'ANOVA à un facteur contrôlé, avec  $p =$  nombre de laboratoires,  $n =$  nombre de répétitions par laboratoire,  $N = np =$  nombre total de déterminations expérimentales,  $\sigma_r^2 =$  variance résiduelle,

$\sigma_T^2 =$  variance totale,  $\sigma_L^2 =$  variance due au facteur laboratoire,  $\bar{x}_i = \frac{1}{n} \sum_{\alpha=1}^n x_{i\alpha}$ ,  $\bar{x} = \frac{1}{N} \sum_{i=1}^p \sum_{\alpha=1}^n x_{i\alpha}$

Source de variations	Somme des carrés Q	Nombre de degrés de liberté : v	Quantité estimée par $q = Q/n$
Facteur Laboratoire	$Q_L = n \sum_{i=1}^p (\bar{x}_i - \bar{x})^2$	$p - 1$	$\sigma_r^2 + n\sigma_L^2$ ( $q_L$ )
Résiduelle	$Q_r = \sum_{i=1}^p \sum_{\alpha=1}^n (x_{i\alpha} - \bar{x}_i)^2$	$N - p$	$\sigma_r^2$ ( $q_r$ )
Totale	$Q_T = \sum_{i=1}^p \sum_{\alpha=1}^n (x_{i\alpha} - \bar{x})^2$	$N - 1$	$\sigma_T^2$

- l'étalonnage (sortie du domaine linéaire, validité dans le temps...);
- l'exploitation des chromatogrammes (mode d'intégration, ligne de base...).

Dès lors qu'une source d'erreur potentielle est identifiée, la mise en œuvre d'une analyse de la variance où ce facteur constitue l'un des facteurs contrôlés permet d'aborder la question de sa mise en évidence et de sa quantification éventuelle. Prenons l'exemple, traité dans notre laboratoire, de l'influence du prélèvement par pesée de l'échantillon pour l'analyse d'un principe actif pharmaceutique par chromatographie en phase liquide (CPL). Ce prélèvement par pesée exacte étant susceptible d'introduire une dispersion supplémentaire dans les résultats d'analyse, la question s'est posée de déterminer à partir de quel niveau ce facteur n'avait plus d'influence significative, de façon à introduire dans le protocole une pesée minimale. Ce problème a été traité en utilisant l'analyse de la variance à un facteur contrôlé, le prélèvement par pesée, dans des conditions opératoires telles que les autres facteurs potentiellement influents (volume de dilution, volume injecté...) ne le soient pas. L'aire des pics chromatographiques étant proportionnelle à la quantité injectée, les résultats ont été exploités sous forme d'aires normalisées par la masse prélevée.

Le tableau II rassemble les résultats pour le plan d'expériences comportant 7 prélèvements et 3 injections par préparation, dans le cas d'une masse voisine de 25 mg.

L'exploitation des résultats du tableau II peut être menée comme suit. En premier lieu, il apparaît que la répétabilité de l'analyse se caractérise par un écart type estimé  $\hat{\sigma}_r = \sqrt{2,93 \cdot 10^6} = 1\,712$ . Ensuite la question de l'influence du facteur prélèvement se pose. Faisons l'hypothèse que celle-ci est nulle, ce qui revient à supposer que  $\sigma_p^2 = 0$ . Dans ce cas, comme le montre la dernière colonne du tableau II,  $q_p$  et  $q_r$  sont deux estimations indépendantes de la même variance  $\sigma_r^2$  et le rapport  $F = \frac{q_p}{q_r}$  correspond à une

variable de Fischer Snedecor à 6 et 14 degrés de liberté. La valeur expérimentale observée pour  $F$  est 3,05. Or la valeur limite qu'une variable de Fischer Snedecor à 6 et 14 degrés de liberté a 95 chances sur 100 de ne pas dépasser est 2,85. Si l'on accepte de se tromper 5 fois sur 100 en rejetant l'hypothèse,

on peut conclure que dans le cas présent l'hypothèse de non-influence du facteur prélèvement peut être rejetée. De manière plus synthétique, avec  $\alpha$  fixé à 5 %, le facteur prélèvement apparaît comme significativement influent. Partant, il est possible de donner une estimation de la variance de prélèvement (3).

$$\hat{\sigma}_p = \sqrt{\frac{q_p - q_r}{3}} = \sqrt{\frac{8,93 \cdot 10^6 - 2,93 \cdot 10^6}{3}} = 1\,410. \quad (3)$$

La même structure expérimentale a été répétée pour d'autres niveaux de masse pesée. Les conclusions sont rassemblées dans le tableau III.

**Tableau III. Influence du facteur prélèvement pour différentes masses pesées.**

Masse pesée (mg)	Influence du facteur prélèvement
10	significative
25	significative
50	non significative
60	non significative

Il apparaît que, dans les conditions de notre analyse, l'influence du facteur prélèvement n'est plus mise en évidence dès lors que la masse pesée est d'au moins 50 mg. Si l'on désire donc que l'étape de prélèvement n'introduise pas de dispersion significative sur le résultat des analyses, le niveau limite de la masse pesée en dessous duquel il ne faut pas descendre est 50 mg.

La technique de l'analyse de la variance peut être étendue à plus d'un facteur contrôlé et autorise le test de l'influence de plusieurs sources d'erreur potentielles. Le dépouillement suivant la même démarche que dans le cas d'un seul facteur permet de quantifier la dispersion apportée par les facteurs dont l'influence est mise en évidence. On peut ainsi quantifier simultanément les effets dispersifs générés à chaque étape d'une analyse et rechercher les conditions où ils peuvent être minimisés.

On conçoit facilement l'existence de cas où l'influence d'un facteur contrôle ne puisse être rendue non significative. Se pose alors la question d'en minimiser les répercussions

**Tableau II. Tableau d'ANOVA expérimental, prélèvement par pesée exacte d'une masse voisine de 25 mg.  $\sigma_p^2$  représente la variance due au facteur prélèvement.**

Source de variations	Q	v	q	Quantité estimée par q	
Prélèvement	5,36.10 <sup>7</sup>	6	8,93.10 <sup>6</sup>	$\sigma_r^2 + 3\sigma_p^2$	( $q_p$ )
Résiduelle	4,11.10 <sup>7</sup>	14	2,93.10 <sup>6</sup>	$\sigma_r^2$	( $q_r$ )
Totale	9,47.10 <sup>7</sup>	20		$\sigma_r^2$	

sur les résultats d'analyse. Prenons l'exemple du facteur préparation des échantillons tel que nous l'avons rencontré en pratique pour la détermination par CPL des constituants principaux d'un antibiotique, la Spiramycine [1-2]. L'analyse de variance pratiquée ayant amené à la conclusion d'une influence significative du facteur laboratoire comme du facteur préparation, les écart-types correspondant  $\widehat{\sigma}_L$  et  $\widehat{\sigma}_{\text{prep}}$  ont pu être estimés. La structure des expériences en

plan pyramidal (chaque laboratoire n'effectue d'injections que pour ses propres préparations) permet de déterminer que la variance de reproductibilité est donnée par la relation (4)

$$\sigma_r^2 = \sigma_L^2 + \frac{\sigma_{\text{prep}}^2}{r} + \frac{\sigma_r^2}{nr} \quad (4)$$

dans laquelle  $r$  désigne le nombre de préparations et  $n$  le nombre d'injections par préparation. Il est facile de vérifier que dans ce cas, et à charge de travail pratiquement identique, il est préférable pour améliorer la reproductibilité, d'effectuer 3 préparations et 2 injections par préparation plutôt que 2 préparations et 3 injections par préparation. On notera qu'en outre le fait de conduire plus de deux préparations permet de détecter une erreur grossière éventuelle.

### Propagation des erreurs aléatoires

Il est fréquent, en analyse chimique, que la détermination d'une grandeur inconnue, par exemple la teneur d'un analyte dans un échantillon, dont on a vu qu'elle ne se conçoit qu'en termes d'estimation, implique la combinaison de plusieurs grandeurs, qui ne seront elles-mêmes qu'estimées avec plus ou moins d'incertitude.

Prenons l'exemple de l'analyse par CPL avec étalonnage externe, où  $\tau_s$ , le titre en analyte d'une solution inconnue, se déduit de la mesure du rapport de l'aire des pics chromatographiques obtenus par injection d'un même volume de la solution inconnue et d'une solution étalon de titre  $\tau_e$ .  $A_s$  et  $A_e$  représentent respectivement les aires obtenues pour la solution inconnue et l'étalon.

$$\tau_s = \tau_e \frac{A_s}{A_e} \quad (5)$$

Pour accéder à un intervalle de confiance, il faut estimer l'écart-type, ce qui amène à traiter le problème de la propagation des erreurs aléatoires, sachant que

- $\tau_e$  n'est pas connu sans erreur, mais fait intervenir un  $\sigma_{\tau_e}$  ;
- $A_s$  et  $A_e$  sont estimés à partir des aires moyennes  $\bar{A}_s$  et  $\bar{A}_e$  observées pour des séries de respectivement  $n_s$  et  $n_e$  injections.

Considérons une fonction  $C$  de deux variables aléatoires  $X_1$  et  $X_2$ .

$$C = f(X_1, X_2). \quad (6)$$

La valeur  $c$  de  $C$  au point de coordonnées  $(x_{1,0}, x_{2,0})$  est estimée à partir des valeurs moyennes  $\bar{x}_1$  et  $\bar{x}_2$  des réalisations de  $X_1$  et  $X_2$ .

$$c = f(\bar{x}_1, \bar{x}_2). \quad (7)$$

Sous réserve que  $X_1$  et  $X_2$  soient indépendantes, on peut à partir du théorème de propagation des erreurs [8] donner l'expression de la variance de  $C$  au point de coordonnées  $(x_{1,0}, x_{2,0})$ .

$$\sigma_C^2(x_{1,0}, x_{2,0}) = \sigma_{x_1}^2 [f'_{x_1}(x_{1,0}, x_{2,0})]^2 + \sigma_{x_2}^2 [f'_{x_2}(x_{1,0}, x_{2,0})]^2 \quad (8)$$

où  $f'_{x_1}$  et  $f'_{x_2}$  représentent respectivement les dérivées de  $f$  par rapport à  $x_1$  et  $x_2$ .

En passant aux variances estimées on obtient :

$$\widehat{\sigma}_C^2 = \widehat{\sigma}_{x_1}^2 [f'_{x_1}(\bar{x}_1, \bar{x}_2)]^2 + \widehat{\sigma}_{x_2}^2 [f'_{x_2}(\bar{x}_1, \bar{x}_2)]^2 \quad (9)$$

Si l'on considère désormais que l'on dispose de  $n_1$  réalisations pour  $X_1$  et de  $n_2$  réalisations pour  $X_2$ .

$$\widehat{\sigma}_C^2 = \frac{\widehat{\sigma}_{x_1}^2}{n_1} [f'_{x_1}(\bar{x}_1, \bar{x}_2)]^2 + \frac{\widehat{\sigma}_{x_2}^2}{n_2} [f'_{x_2}(\bar{x}_1, \bar{x}_2)]^2 \quad (10)$$

En appliquant cette formule au cas de l'analyse CPL avec étalonnage externe décrit précédemment et dont l'expression de la fonction  $f$  est donnée équation (5) on obtient :

$$\widehat{\sigma}_{\tau_s} = \sqrt{\widehat{\sigma}_{\tau_e}^2 \left(\frac{\bar{A}_s}{\bar{A}_e}\right)^2 + \frac{\widehat{\sigma}_{A_s}^2}{n_s} \left(\frac{\tau_e}{\bar{A}_e}\right)^2 + \frac{\widehat{\sigma}_{A_e}^2}{n_e} \left(-\frac{\tau_e \bar{A}_s}{\bar{A}_e^2}\right)^2} \quad (11)$$

Si l'on remarque qu'il est légitime de considérer que  $\sigma_{A_s}^2 = \sigma_{A_e}^2 = \sigma_r^2$ , variance de répétabilité estimée par, on peut écrire :

$$\widehat{\sigma}_{\tau_s} = \sqrt{\widehat{\sigma}_{\tau_e}^2 \left(\frac{\bar{A}_s}{\bar{A}_e}\right)^2 + \widehat{\sigma}_r^2 \left(\frac{\tau_e}{\bar{A}_e}\right)^2 \left(\frac{1}{n_e} + \frac{1}{n_s} \left(\frac{\bar{A}_s}{\bar{A}_e}\right)^2\right)} \quad (12)$$

Comme généralement on opère dans des conditions telles que  $A_s \cong A_e$  la formule précédente peut se simplifier en

$$\widehat{\sigma}_{\tau_s} \cong \sqrt{\widehat{\sigma}_{\tau_e}^2 + \widehat{\sigma}_r^2 \left(\frac{\tau_e}{\bar{A}_e}\right)^2 \left(\frac{1}{n_e} + \frac{1}{n_s}\right)} \quad (13)$$

Remarque : Avec l'hypothèse que les titres de l'étalon et du soluté sont proches, et en utilisant un développement limité d'ordre 1, il est possible de donner l'expression du coefficient de variations du titre du soluté  $\widehat{CV}_{\tau_s}$  défini comme

$\frac{\hat{\sigma}_{\tau_s}}{\tau_s}$ , en fonction du coefficient de variations du titre de l'étalon  $CV_{\tau_e}$  et du coefficient de variations de répétabilité de l'analyse  $CV_r$ .

$$CV_{\tau_s} = CV_{\tau_e} + \frac{1}{2} CV_r \left( \frac{1}{n_s} + \frac{1}{n_e} \right). \quad (14)$$

### Détection des erreurs systématiques

Un test de la justesse d'une méthode d'analyse ou d'une chaîne métrologique peut être construit directement si l'on dispose de produits de référence : étalons ou matériaux certifiés.

#### Étude de la justesse lorsque l'on dispose de produits de référence

Pour illustrer le principe d'un tel test et les risques d'erreur dans les conclusions adoptées, considérons le cas d'une méthode d'analyse dont l'application à une solution de concentration connue  $c_e$ , a conduit à l'issue de  $n$  répétitions à une concentration moyenne observée. Supposons, de plus, que l'écart type caractéristique de la population des résultats soit connu ou, parce qu'estimé avec un grand nombre de degrés de liberté, puisse être considéré comme tel. La

moyenne  $\bar{c}$  est, si la méthode est juste et seulement si elle l'est, la réalisation d'une variable aléatoire  $\bar{C}$  distribuée suivant une loi normale d'espérance  $c_e$  et d'écart type  $\frac{\sigma}{\sqrt{n}}$ .

On peut alors définir un intervalle à l'intérieur duquel  $\bar{C}$  a une probabilité donnée, par exemple 95 chances sur 100, de prendre sa réalisation. Si la valeur obtenue n'appartient pas à cet intervalle, comme il s'agit d'une valeur appartenant au groupe de celles observées moins de 5 fois sur 100, donc

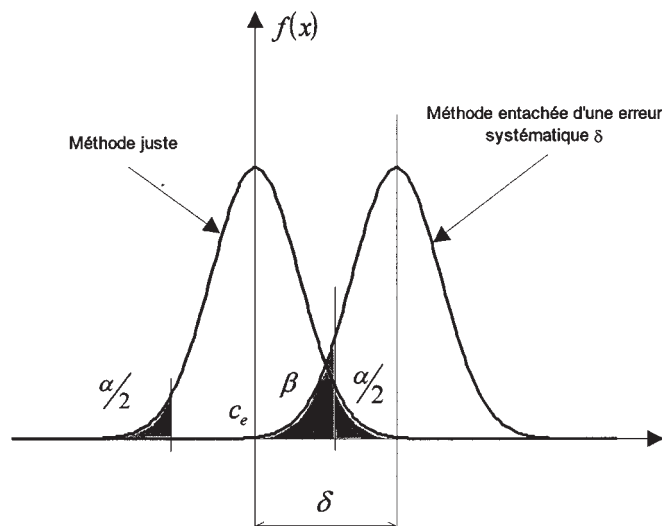


Figure 5. Risques de première et de seconde espèce.

d'apparition très peu probable dans le cadre de l'hypothèse faite, on est conduit à remettre celle-ci en cause : on conclut au rejet de l'hypothèse de justesse de la méthode. Mais il existe un risque égal à  $\alpha$  (= 5 chances sur 100) pour que cette décision ne soit pas justifiée. Supposons maintenant qu'en réalité la méthode présente une erreur systématique égale à  $\delta$ . La règle de décision précédente fait qu'il existe un risque de ne pas la détecter à l'issue du test effectué, correspondant à la probabilité  $\beta$  (Fig. 5). Les deux risques  $\alpha$  et  $\beta$  définissent respectivement les risques dits de première et de seconde espèce de conclusions inadéquates (Tab. IV).

La règle de décision est généralement construite à partir d'un risque de première espèce fixé *a priori*. Le risque de seconde espèce est ensuite évalué dans un cas particulier, correspondant par exemple au risque de ne pas détecter une erreur systématique en réalité égale à l'écart type, soit  $\delta = \sigma$ . Si ce risque est jugé trop élevé, comme les deux risques varient de manière opposée, la seule option pour diminuer  $\beta$  sans pour autant augmenter  $\alpha$  consiste à augmenter le nombre de répétitions.

Tableau IV. Risques d'erreurs de première et de seconde espèce dans la conduite d'un test de justesse.

		Réalité (qui est et restera inconnue)	
		Méthode juste	Méthode entachée d'une erreur systématique
Conclusion adoptée à l'issue du test	absence d'erreur systématique	décision correcte Probabilité : $(1 - \alpha)$	décision incorrecte Probabilité : $\beta$
	présence d'une erreur systématique	décision incorrecte Probabilité : $\alpha$	décision correcte Probabilité : $(1 - \beta)$

N.B. : Si l'écart type  $\sigma$  n'est pas connu, mais seulement estimé avec un nombre de degrés de liberté  $\nu$ , la construction du test s'effectue en substituant à la loi normale une loi de Student comportant  $\nu$  degrés de liberté.

Cependant, le chimiste analyste peut être amené à s'interroger sur la détection d'une erreur systématique éventuelle, non pas seulement à un niveau de concentration donné de l'analyte, mais sur toute la gamme d'utilisation de la méthode, d'autant qu'elle peut y être variable. L'approche statistique consiste alors à appliquer la méthode à différentes solutions étalons réparties sur la gamme en question. Si la méthode est juste et seulement dans ce cas, la courbe représentant les variations de l'espérance de chaque population de résultats possibles, définie chacune par une concentration de référence, en fonction de la valeur vraie de la concentration de l'analyte, doit se confondre avec la première bissectrice. Si c'est une droite parallèle à cette bissectrice, on a l'indice d'une erreur systématique constante ; si c'est une droite de pente différente de l'unité, on a l'indice d'une erreur systématique proportionnelle à la concentration, due par exemple à une opération de séparation de l'échantillon (Fig. 6). Dans la pratique, l'espérance n'est pas accessible expérimentalement : elle n'est qu'estimée, au mieux, par la moyenne des résultats de mesures répétées pour chaque concentration de référence. La loi de variation précédente doit être établie en utilisant une méthode de régression, telle la régression linéaire simple ou pondérée [12]. L'application de tests statistiques permet alors de juger de l'adéquation du modèle linéaire, et en cas de réponse positive, la comparaison de l'ordonnée à l'origine et de la pente aux caractéristiques de la bissectrice des axes, à savoir respectivement 0 et 1.

On notera que l'utilisation de matériaux de référence certifiés en lieu et place de solutions étalons de l'analyte ne contenant que celui-ci permet la recherche d'erreurs systématiques résultant d'effets de matrice.

### Étude de la justesse en l'absence de produits de référence

En l'absence de produits de référence, la loi de variation précédente peut être obtenue indirectement au moyen de deux méthodes complémentaires, la méthode des ajouts dosés et la méthode des prises d'essai variables [8].

#### Méthode des ajouts dosés

Des quantités connues de l'analyte sont ajoutées à des prises d'essai identiques de l'échantillon inconnu. Si  $\bar{c}_0$  est la moyenne des résultats obtenus avant ajout,  $\bar{c}_i$  celle observée après un ajout de la quantité  $q_i$ , il est clair que l'étude de la courbe  $(\bar{c}_i - \bar{c}_0) = f(q_i)$  permet de reconstituer

la courbe de variation de l'espérance en fonction de la concentration vraie à une translation inconnue près, correspondant à la teneur initiale. Par conséquent, si la courbe est une droite dont la pente n'est pas significativement différente de l'unité, la méthode des ajouts dosés permet de conclure à l'absence d'erreur systématique variable. En revanche, cette méthode se révèle totalement impuissante en ce qui concerne la détection des erreurs systématiques constantes.

#### Méthode des prises d'essai variables

La méthode d'analyse est appliquée à des prises d'essai différentes de l'échantillon initial. Supposons, pour exposer le principe de la méthode, que l'on opère sur une prise d'essai double de celle qui a produit le résultat  $\bar{c}_0$ . Si la moyenne obtenue alors est  $\bar{c}'$ , l'abscisse du point d'ordonnée  $(\bar{c}' - \bar{c}_0)$

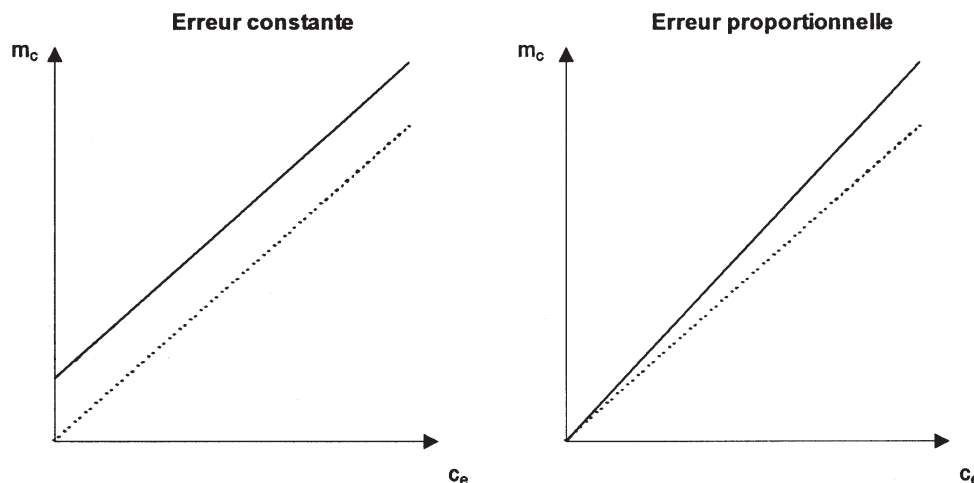


Figure 6. Erreur constante et erreur proportionnelle.



sur la courbe précédente fournit une estimation de  $c_0$ , que l'on peut comparer à  $\bar{c}_0$ . En généralisant au cas d'une prise d'essai dans un rapport  $k$  avec la première, l'abscisse déterminée sur la courbe fournit une estimation de  $(k - 1)c_0$ . En faisant varier  $k$ , on peut donc examiner l'homogénéité des valeurs estimées pour  $c_0$ . La méthode des ajouts dosés permet donc de détecter une erreur systématique constante non accessible par la première méthode.

### Conclusion

Les méthodes statistiques que l'on peut définir comme le moyen le plus objectif de traiter et d'exploiter l'information disponible, peuvent amener un appui décisif dans la maîtrise de la qualité des analyses. Elles fournissent les outils nécessaires à une optimisation raisonnée des conditions opératoires, en permettant l'identification des causes d'erreur, puis leur quantification et partant, leur élimination ou à tout le moins une diminution de leur effet, lorsqu'elles sont inévitables. C'est cet éclairage que nous avons voulu apporter, en essayant de rester simple, sans pour autant sacrifier à la rigueur nécessaire. Il existe de nombreux prolongements à l'approche présentée qui peuvent apporter une aide efficace à l'analyste, d'autant que la commercialisation de logiciels dédiés permet désormais de franchir l'écueil de l'aspect calculatoire. Citons, par exemple, la mise en œuvre de plans d'expériences pour l'étude de la robustesse, laquelle caractérise l'insensibilité relative d'une méthode à de petites variations involontaires des paramètres expérimentaux [13].

### Références

1. Jardy, A.; Vial, J.; Ménier, I. *Analisis* **1997**, 25, 106-111.
2. Vial, J.; Ménier, I.; Jardy, A.; Anger, P.; Brun, A.; Burbaud, L. *J. Chromatogr. B* **1998**, 708, 131-143.
3. Martin-Esteban, A.; Fernandez, P.; Camara, C.; Kramer, G. N.; Maier, E. A. *Intern. J. Anal. Chem.* **1997**, 67, 125-141.
4. Terms and Definitions Used in Connection with Reference Materials (ISO Guide 30-1981) International Organization for Standardization, Genève, 1981.
5. Wise, S. A. In: *Environmental Analysis*; Barcelo, D. Ed., Elsevier: Amsterdam, 1993, Chap. 12.
6. Maier, E. A. In: *Environmental Analysis*; Barcelo, D. Ed., Elsevier: Amsterdam, 1993, Chap. 11.
7. Massart, D. L.; Vandeginste, B. G. M.; Deming, S. N.; Michotte, Y.; Kaufman, L. In: *Chemometrics: A textbook*; Elsevier: Amsterdam, 1990.
8. Commissariat à l'énergie atomique, In: *Statistique appliquée à l'exploitation des mesures Tome I & Tome II*; Masson: Paris, 1978.
9. Ceresta, In: *Aide-mémoire pratique des techniques statistiques*; Ceresta: Paris, 1994.
10. Barcelo, D.; Hennion, M. C. In: *Trace determination of pesticides and their degradation products in water*; Elsevier: Amsterdam, 1997, Ch. 2, pp 95-155.
11. Funk, W.; Dammann, V.; Donnevert, G. In: *Quality Assurance in Analytical Chemistry*; VCH: Weinheim, 1995, Ch. 3, pp 140-152.
12. Miller, J. C.; Miller, J. N. In: *Statistics for Analytical Chemistry*; Second Edition, Ellis Horwood: Chichester, 1988, Ch. 5, pp 101-136.
13. Hendriks, M. M. W. B.; De Boer, J. H.; Smilde, A. K. In: *Robustness of analytical chemical methods and pharmaceutical technological products*; Elsevier: Amsterdam, 1996.