

ANALYSE STATISTIQUE DES DONNÉES EXPÉRIMENTALES

Konstantin PROTASSOV



17, avenue du Hoggar
Parc d'Activité de Courtabœuf, BP 112
91944 Les Ulis Cedex A, France

Grenoble Sciences

Grenoble Sciences poursuit un triple objectif :

- réaliser des ouvrages correspondant à un projet clairement défini, sans contrainte de mode ou de programme,
- garantir les qualités scientifique et pédagogique des ouvrages retenus,
- proposer des ouvrages à un prix accessible au public le plus large possible.

Chaque projet est sélectionné au niveau de Grenoble Sciences avec le concours de referees anonymes. Puis les auteurs travaillent pendant une année (en moyenne) avec les membres d'un comité de lecture interactif, dont les noms apparaissent au début de l'ouvrage. Celui-ci est ensuite publié chez l'éditeur le plus adapté.

(Contact : Tél. : (33)4 76 51 46 95 - E-mail : Grenoble.Sciences@ujf-grenoble.fr)

Deux collections existent chez EDP Sciences :

- la *Collection Grenoble Sciences*, connue pour son originalité de projets et sa qualité
- *Grenoble Sciences - Rencontres Scientifiques*, collection présentant des thèmes de recherche d'actualité, traités par des scientifiques de premier plan issus de disciplines différentes.

Directeur scientifique de Grenoble Sciences

Jean BORNAREL, Professeur à l'Université Joseph Fourier, Grenoble 1

Comité de lecture pour "Analyse statistique des données expérimentales"

- ◆ J.P. BERTRANDIAS, Professeur à l'Université Joseph Fourier, Grenoble 1
- ◆ C. FURGET, Maître de conférences à l'Université Joseph Fourier, Grenoble 1
- ◆ B. HOUCHMANDZADEH, Directeur de recherches au CNRS, Grenoble
- ◆ M. LESIEUR, Professeur à l'Institut National Polytechnique, Grenoble
- ◆ C. MISBAH, Directeur de recherches au CNRS, Grenoble
- ◆ J.L. PORTESEIL, Professeur à l'Université Joseph Fourier, Grenoble 1
- ◆ P. VILLEMMAIN, Maître de conférences à l'Université Joseph Fourier, Grenoble 1

Grenoble Sciences reçoit le soutien
du **Ministère de l'Éducation nationale**, du **Ministère de la Recherche**,
de la **Région Rhône-Alpes**, du **Conseil général de l'Isère**
et de la **Ville de Grenoble**.

ISBN 2-86883-590-2

© EDP Sciences, 2002

EXTRAITS

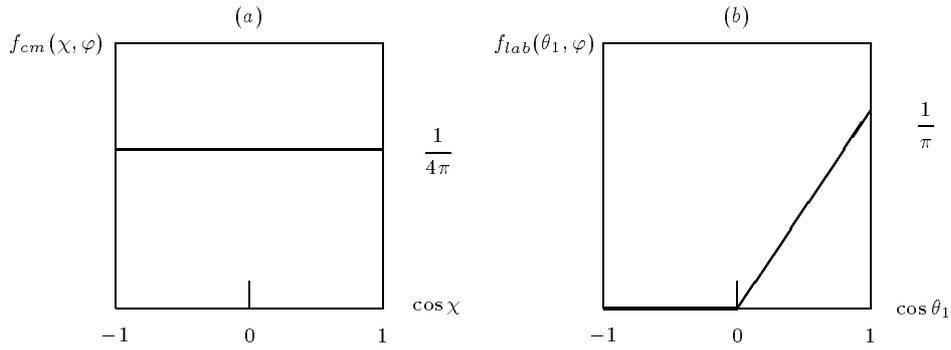


Figure 2.5 : Les distributions angulaires dans le système du centre de masse (a) et dans le système du laboratoire (b)

2.2.4 PRÉCISION DE LA FORMULE DE PROPAGATION DES ERREURS

Nous avons déjà souligné que la formule de propagation des erreurs, largement utilisée dans le traitement des résultats expérimentaux, est une formule approchée (sauf dans le cas presque trivial d'une fonction linéaire). Cette approximation est parfois assez grossière puisque pour obtenir la formule de propagation des erreurs nous avons utilisé la relation (49) : $y(x) \simeq y(\bar{x})$, alors que toute la statistique est basée, par la définition de la variance, sur l'importance de la différence entre $\bar{y} = \overline{x^2}$ et $\bar{y} \simeq \bar{x}^2$.

Dans certains cas, nous pouvons obtenir l'expression exacte de la variance σ_y^2 sans utiliser la formule de propagation des erreurs. Considérons l'exemple très simple d'une fonction produit de deux variables indépendantes :

$$y = x_1 x_2.$$

Cette fonction peut être mise sous la forme équivalente :

$$y = \mu_1 \mu_2 + \mu_1(x_2 - \mu_2) + \mu_2(x_1 - \mu_1) + (x_1 - \mu_1)(x_2 - \mu_2), \quad (80)$$

c'est-à-dire sous la forme d'un développement en série de Taylor au voisinage du point $x_1 = \mu_1, x_2 = \mu_2$. L'expression (80) contient un nombre fini de termes : une constante $\mu_1 \mu_2$; les contributions avec les dérivées premières

$$\left. \frac{\partial y}{\partial x_1} \right|_{x_1=\mu_1, x_2=\mu_2} = \mu_2, \quad \left. \frac{\partial y}{\partial x_2} \right|_{x_1=\mu_1, x_2=\mu_2} = \mu_1 ;$$

et un seul terme avec les dérivées secondes puisque

$$\frac{\partial^2 y}{\partial x_1^2} = \frac{\partial^2 y}{\partial x_2^2} = 0, \quad \frac{\partial^2 y}{\partial x_1 \partial x_2} = 1.$$

Compte tenu de l'indépendance de x_1 et x_2 , nous pouvons calculer exactement la variance de y :

$$\begin{aligned} \sigma_y^2 &\equiv \overline{(y - \mu_1 \mu_2)^2} = \overline{[\mu_1(x_2 - \mu_2) + \mu_2(x_1 - \mu_1) + (x_1 - \mu_1)(x_2 - \mu_2)]^2} \\ &= \mu_1^2 \sigma_{x_2}^2 + \mu_2^2 \sigma_{x_1}^2 + \sigma_{x_1}^2 \sigma_{x_2}^2. \end{aligned}$$

La formule de propagation des erreurs (57)

$$\sigma_y^2 \simeq \mu_1^2 \sigma_{x_2}^2 + \mu_2^2 \sigma_{x_1}^2$$

est obtenue en négligeant le dernier terme dans le développement (80). Ainsi cette formule conduit à une erreur supplémentaire dans le calcul de $(\Delta y)^2 \equiv \sigma_y^2$ égale à $\sigma_{x_1}^2 \sigma_{x_2}^2$.

On pourrait penser qu'il est facile d'améliorer la formule de propagation des erreurs en poussant plus loin le développement de la fonction en série de Taylor. Cette proposition apparaît dans certains livres sur l'analyse des données. Techniquement, c'est un exercice simple, bien qu'il soit assez pénible (il faut faire très attention et garder correctement tous les termes de même ordre dans le développement et dans les calculs intermédiaires). Cependant des problèmes majeurs apparaissent dans cette voie.

Considérons l'exemple simple d'une fonction d'une seule variable $y = y(x)$. Comme pour la formule de propagation des erreurs, développons cette fonction en série de Taylor au voisinage de $x = \mu_x \equiv \bar{x}$:

$$y(x) \simeq y(\bar{x}) + y'(\bar{x})(x - \bar{x}) + \frac{1}{2!}y''(\bar{x})(x - \bar{x})^2 + \frac{1}{3!}y'''(\bar{x})(x - \bar{x})^3 + \dots$$

Nous conservons volontairement le terme du troisième ordre car il donnera en fait une contribution à la variance du même ordre que le terme du second ordre. La valeur moyenne de \bar{y} prend alors la forme

$$\bar{y} \simeq y(\bar{x}) + \frac{1}{2!}y''(\bar{x})\sigma_x^2 + \frac{1}{3!}y'''(\bar{x})\mu_{x3},$$

où apparaît le troisième moment de la distribution $\mu_{x3} \equiv \overline{(x - \bar{x})^3}$ introduit en (12), qui caractérise l'asymétrie de la distribution de x . Ainsi, pour la variance, nous obtenons

$$\sigma_y^2 \equiv \overline{(y - \bar{y})^2} \simeq y'^2 \sigma_x^2 + y' y'' \mu_{x3} + \frac{1}{3} y' y''' \mu_{x4} + \frac{1}{4} y''^2 (\mu_{x4} - \sigma_x^4),$$

où est en outre introduit le quatrième moment $\mu_{x4} \equiv \overline{(x - \bar{x})^4}$.

Le problème est résolu formellement mais le prix à payer est l'introduction de moments centraux d'ordres supérieurs non utilisés jusqu'à présent et dont la détermination expérimentale peut s'avérer délicate. Pour obtenir une expression plus précise de la variance, on a sacrifié la simplicité de la description des grandeurs physiques. Rappelons, que dans la plupart des situations, nous travaillons avec des distributions gaussiennes. La prise en compte du terme linéaire dans la formule de propagation des erreurs nous garantit la conservation du langage utilisé (la variable y est aussi décrite par la distribution normale). Il est vrai que, si x est décrite par une distribution gaussienne, nous pouvons exprimer tous les moments d'ordres supérieurs à l'aide de la variance (voir (27)), mais le problème vient du fait que la variable y n'est plus gaussienne (on peut vérifier que la distribution de y est asymétrique : $\mu_{y3} \neq 0$). Quand la distribution de y est gaussienne, un écart-type σ_y a une interprétation précise. Dans le cas contraire il peut la perdre. La question qui se pose est de savoir s'il

est l'intéressant d'obtenir une expression plus précise de l'incertitude d'une grandeur physique si l'on ne peut plus l'interpréter avec précision.

Pour mieux comprendre, étudions sur un exemple le "passage" d'une distribution gaussienne à une distribution plus complexe. Soient x_1 et x_2 deux variables gaussiennes. Quelle est la distribution de leur rapport

$$y = \frac{x_1}{x_2} ?$$

Appliquons l'approche générale présentée dans le paragraphe 2.2.2. Il faut passer des variables x_1 et x_2 aux variables y et $z = x_2$ (cette dernière joue le rôle d'une variable auxiliaire) et intégrer sur z .

Pour simplifier les relations, supposons que les valeurs moyennes μ_i sont positives et que les incertitudes sont faibles par rapport aux valeurs moyennes ($\sigma_i \ll \mu_i$). Cela signifie que la distribution cherchée reste proche d'une distribution gaussienne. Si $f(x_1)$ et $f(x_2)$ sont les fonctions de distribution des variables x_1 et x_2

$$f(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left\{ -\frac{(x_i - \mu_i)^2}{2\sigma_i^2} \right\},$$

selon (77), la fonction de distribution $g(y)$ de la variable y prend la forme

$$g(y) = \int_{-\infty}^{+\infty} f(x_1(y, z))f(x_2(y, z)) \left| \frac{\partial(x_1, x_2)}{\partial(y, z)} \right| dz.$$

Le Jacobien de la transformation $x_1 = yz$, $x_2 = z$ est égal à

$$\left| \frac{\partial(x_1, x_2)}{\partial(y, z)} \right| = \begin{vmatrix} \frac{\partial x_1}{\partial y} & \frac{\partial x_2}{\partial y} \\ \frac{\partial x_1}{\partial z} & \frac{\partial x_2}{\partial z} \end{vmatrix} = \begin{vmatrix} z & 0 \\ y & 1 \end{vmatrix} = z = x_2.$$

Ainsi l'intégrale $g(y)$ prend la forme

$$g(y) = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp \left\{ -\frac{(yz - \mu_1)^2}{2\sigma_1^2} - \frac{(z - \mu_2)^2}{2\sigma_2^2} \right\} z dz.$$

Cette dernière intégrale peut être calculée si l'on utilise la valeur de l'intégrale auxiliaire²

$$J(A, B) = \int_{-\infty}^{+\infty} \exp \{-Ax^2 + Bx\} x dx = \frac{\sqrt{\pi}}{2} \frac{B}{A^{3/2}} \exp \left\{ \frac{B^2}{4A} \right\}.$$

2 L'astuce pour calculer $J(A, B)$ est classique : il faut utiliser la méthode de dérivation par rapport au paramètre B :

$$\int_{-\infty}^{+\infty} e^{-Az^2+Bz} z dz = \frac{\partial}{\partial B} \int_{-\infty}^{+\infty} e^{-Az^2+Bz} dz = \sqrt{\frac{\pi}{A}} \frac{\partial}{\partial B} e^{B^2/4A}.$$

La dernière intégrale se remène à l'intégrale connue (25) par le changement linéaire de variable $y = \sqrt{A}z - B/2\sqrt{A}$.

Avec $A = \frac{y^2}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}$ et $B = y\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}$, on trouve finalement après quelques calculs laborieux mais sans difficulté majeure

$$g(y) = \frac{1}{\sqrt{2\pi}\Delta(y)} \frac{\Delta_0^2(y)}{\Delta^2(y)} \exp\left\{-\frac{(y-y_0)^2}{2\Delta^2(y)}\right\} \approx \frac{1}{\sqrt{2\pi}\Delta(y)} \exp\left\{-\frac{(y-y_0)^2}{2\Delta^2(y)}\right\}. \quad (81)$$

Dans cette expression

$$y_0 = \frac{\mu_1}{\mu_2}, \quad \Delta^2(y) = \frac{\sigma_1^2}{\mu_2^2} + y^2 \frac{\sigma_2^2}{\mu_2^2}, \quad \Delta_0^2(y) = \frac{\sigma_1^2}{\mu_2^2} + yy_0 \frac{\sigma_2^2}{\mu_2^2}.$$

La fonction (81) s'écrit sous une forme qui ressemble beaucoup (surtout si l'on fait l'approximation supplémentaire $\Delta_0^2(y)/\Delta^2(y) \approx 1$) à la distribution de Gauss, mais sa largeur dépend de y .

Un exemple d'une telle distribution est tracé sur la figure 2.6 (pour $\mu_1/\mu_2 = 1$, $\sigma_1/\mu_1 = 0,3$ et $\sigma_2/\mu_2 = 0,2$).

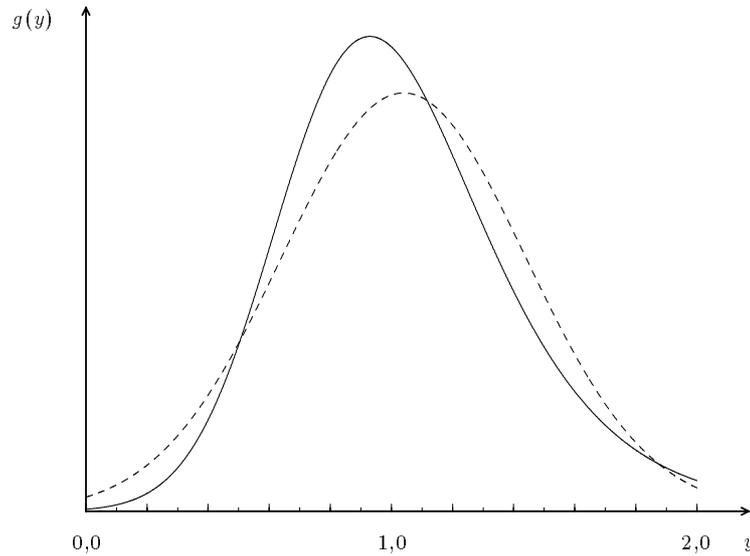


Figure 2.6 : La fonction de distribution $g(y)$ de $y = x_1/x_2$ (ligne continue) comparée à une fonction gaussienne (ligne pointillée).

On constate que, lorsque les incertitudes relatives sont faibles ($\sigma_i \ll \mu_i$), la fonction de distribution $g(y)$ est très proche d'une gaussienne : c'est une fonction qui est très piquée au voisinage de $y = y_0 = \mu_1/\mu_2$ (on peut donc garder la dépendance rapide de y dans la fonction exponentielle, mais remplacer partout ailleurs y par y_0) avec une largeur σ_y dont le carré est égal à

$$\sigma_y^2 \approx \Delta^2(y)|_{y=y_0} = \frac{\sigma_1^2}{\mu_2^2} + \frac{\mu_1^2 \sigma_2^2}{\mu_2^2 \mu_2^2} = \left(\frac{\mu_1}{\mu_2}\right)^2 \left(\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2}\right).$$

Donc, en première approximation, on retrouve une distribution gaussienne avec une moyenne $y_0 = \mu_1/\mu_2$ et une incertitude σ_y en parfait accord avec la formule de propagation des erreurs (55).

Si l'on veut ne pas se limiter à de cette approximation, on peut remarquer que la fonction $g(y)$ n'est pas tout à fait symétrique par rapport à $y = y_0$ et aucune gaussienne, même avec une largeur calculée à partir de la formule de propagation des erreurs améliorée, ne peut décrire correctement cette distribution. Ce fait est illustré sur la figure 2.6 où la fonction de distribution (81) est comparée avec une fonction gaussienne pour laquelle la moyenne $\overline{y_{\text{sup}}}$ et la variance σ_{sup}^2 sont calculées à l'ordre supérieur du développement en série de Taylor³

$$\begin{aligned}\overline{y_{\text{sup}}} &\approx \frac{\mu_1}{\mu_2} \left(1 + \frac{\sigma_2^2}{\mu_2^2} \right) \approx 1,04, \\ \sigma_{\text{sup}}^2 &\approx \left(\frac{\mu_1}{\mu_2} \right)^2 \left[\frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} + 3 \frac{\sigma_1^2 \sigma_2^2}{\mu_1^2 \mu_2^2} + 8 \frac{\sigma_2^4}{\mu_2^4} \right] \approx 0,15.\end{aligned}$$

Notons que ces valeurs sont très proches de la moyenne μ_y et de la variance σ_y^2 calculées avec la fonction de distribution (81) :

$$\mu_y \approx 1,05, \quad \sigma_y^2 \approx 0,16.$$

Néanmoins, la différence entre ces deux fonctions est évidente.

On remarquera que la nouvelle fonction (81) dépend de trois variables $y_0 = \mu_1/\mu_2$, σ_1/μ_1 et σ_2/μ_2 , tandis qu'une gaussienne ne dépend que de deux variables. En principe, des mesures précises de la fonction de distribution $g(y)$ peuvent permettre d'avoir non seulement des informations sur la variable y mais aussi sur x_1 et x_2 (une des quatre caractéristiques des distributions initiales $\mu_1, \sigma_1, \mu_2, \sigma_2$ restera toujours inconnue mais on pourra avoir les rapports entre elle et les autres).

En conclusion de ce paragraphe, on constate que "l'amélioration" de la formule de propagation des erreurs, grâce à l'augmentation du nombre de termes dans le développement en série de Taylor, ne représente aucune difficulté. Mais cela n'a pas beaucoup d'intérêt puisque l'interprétation du résultat obtenu en termes de probabilités reste assez limité.

2.3 NIVEAU DE CONFIANCE ET INTERVALLE DE CONFIANCE

Nous avons déjà étudié des distributions très différentes : symétriques et asymétriques ; définies sur un intervalle fini, demi-infini et infini ; déterminées par un ou plusieurs paramètres. Si nous conservons la même approche, la description des données expérimentales devient assez lourde (pour chaque grandeur physique on est obligé d'indiquer la loi de probabilité et ses paramètres). Sans doute, une telle approche est indispensable pour rester précis dans la description des données (sans approximer les distributions de toutes les grandeurs par une loi gaussienne). Cependant, il est possible de

³ Nous laissons au lecteur le soin de retrouver ces expressions.

3.2 DISTRIBUTION DE STUDENT

Pour pouvoir interpréter les résultats expérimentaux en termes de m (82) et de s_m (88), on a besoin de la fonction de distribution de la variable

$$t \equiv y(x_1, x_2, \dots, x_n) = \frac{m - \mu}{s_m} \quad (106)$$

où m et s_m sont définies par (82) et (88).

La solution du problème est relativement simple si nous exprimons cette fonction sous la forme

$$t = \sqrt{n-1} \frac{\frac{m-\mu}{\sigma/\sqrt{n}}}{\sqrt{(n-1)n \cdot s_m^2/\sigma^2}} \equiv \frac{y_1}{\sqrt{y_2}} \cdot \sqrt{n-1},$$

où

$$y_1 = \frac{m - \mu}{\sigma/\sqrt{n}} \quad \text{et} \quad y_2 = (n-1)n \frac{s_m^2}{\sigma^2} = (n-1) \frac{s^2}{\sigma^2}.$$

La variable y_1 a une distribution normale (car tous les x_i ont la même distribution normale) avec la moyenne nulle (83) et la variance unité (84). La variable y_2 est distribuée selon χ_{n-1}^2 comme nous venons de le démontrer (104). Ainsi nous connaissons les distributions de y_1 et de y_2 et nous voulons trouver la distribution du rapport $t = y_1/\sqrt{y_2}$ en utilisant les règles connues de transformation des distributions.

La densité de probabilité de y_1 et y_2 est égale à :

$$g(y_1, y_2) = \frac{e^{-y_1^2/2}}{\sqrt{2\pi}} \cdot \frac{y_2^{\frac{n-1}{2}-1} \cdot e^{-y_2/2}}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}}$$

avec y_1 qui varie de $-\infty$ jusqu'à $+\infty$ et y_2 qui varie de 0 jusqu'à $+\infty$. Transformons d'abord cette densité en faisant le changement de variables

$$\begin{cases} z_1(y_1, y_2) = \frac{y_1}{\sqrt{y_2}} \equiv \frac{t}{\sqrt{n-1}}, \\ z_2(y_1, y_2) = y_2, \end{cases}$$

soit par transformation inverse

$$\begin{cases} y_1(z_1, z_2) = z_1 \sqrt{z_2}, \\ y_2(z_1, z_2) = z_2. \end{cases}$$

Le module du Jacobien de cette transformation est égal à $\sqrt{z_2}$ et, conformément à (77), la nouvelle densité de probabilité $h(z_1, z_2)$ est

$$h(z_1, z_2) = \sqrt{z_2} \cdot \frac{e^{-(z_1^2 z_2)/2}}{\sqrt{2\pi}} \cdot \frac{z_2^{\frac{n-1}{2}-1} \cdot e^{-z_2/2}}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}}.$$

Pour obtenir la densité de probabilité $f(t)$ nous intégrons $h(z_1, z_2)$ par rapport à z_2 et utilisons la relation $f(t) = f(z_1)|dz_1/dt|$:

$$\begin{aligned} f(t) &= \frac{f(z_1)}{\sqrt{n-1}} = \int_0^\infty \frac{h(z_1, z_2)}{\sqrt{n-1}} dz_2 \\ &= \int_0^\infty \sqrt{z_2} \cdot \frac{e^{-(z_1^2 z_2)/2}}{\sqrt{2\pi(n-1)}} \cdot \frac{z_2^{\frac{n-1}{2}-1} \cdot e^{-z_2/2}}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}} dz_2. \end{aligned}$$

Le changement de variable

$$u = \frac{z_2(z_1^2 + 1)}{2}$$

ramène cette intégrale à une fonction Γ .

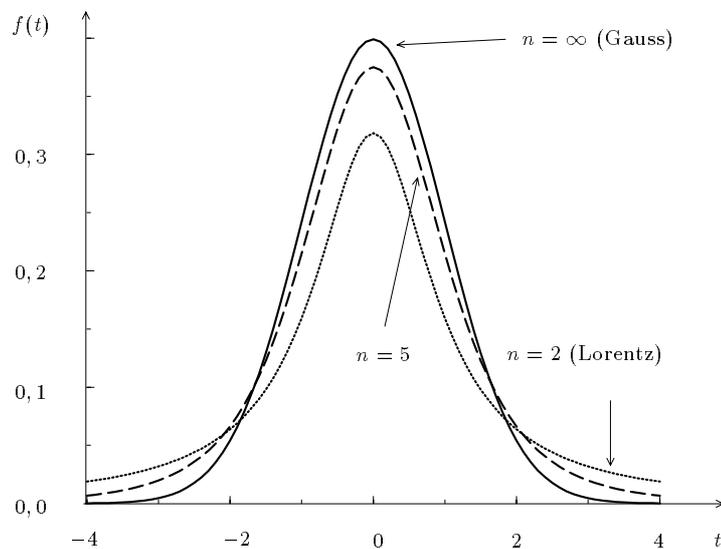


Figure 3.3 : La distribution de Student pour $n = 2$ (distribution de Lorentz), $n = 5$, et $n = \infty$ (distribution de Gauss)

Finalement la distribution $f(t)$ s'écrit

$$f_{n-1}(t) = \frac{C}{\left(1 + \frac{t^2}{n-1}\right)^{n/2}}, \quad (107)$$

où t a $n - 1$ degrés de liberté. Les variables initiales y_1 et y_2 (soit χ_{n-1}^2) en ont 1 et $n - 1$ respectivement. L'intégration sur z_2 a éliminé une variable (un degré de liberté) : $1 + (n - 1) - 1 = n - 1$. La constante C dans l'expression (107) est égale à

$$C = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\Gamma\left(\frac{n-1}{2}\right)}.$$

Pour n donné, les fonctions Γ dans la formule ci-dessus peuvent être explicitées à l'aide de (43) et (44).

Cette fonction (107) est relativement simple. Pour $n = 2$, on retrouve la distribution de Lorentz. Pour $n > 2$, la distribution t de Student représente, grosso modo, une certaine puissance de cette distribution. Vu la discussion du paragraphe 1.3.3, nous pouvons tout de suite dire que, pour n donné, seuls les moments μ_k avec $k < n - 1$ peuvent être définis.

On peut aussi calculer facilement la valeur moyenne et la variance de cette distribution lorsque cette dernière existe :

$$\bar{t} = 0 \quad \text{et} \quad D(t) = \frac{n-1}{n-3}.$$

Dans la limite $n \rightarrow \infty$, la distribution de Student se transforme en distribution gaussienne. La démonstration est simple et peut être réalisée par le lecteur intéressé.

Plusieurs exemples de la distribution de Student sont présentés sur la figure 3.3.

La figure 3.4 est une version élargie de la figure 1.12 ; elle montre les relations qui existent entre les différentes distributions. Notons que nous avons regroupé la distribution Γ (45) et celle de Poisson (36) par suite de la ressemblance formelle de leurs dépendances fonctionnelles. Néanmoins, il ne faut pas oublier que les rôles des variables et des paramètres sont inversés dans ces distributions.

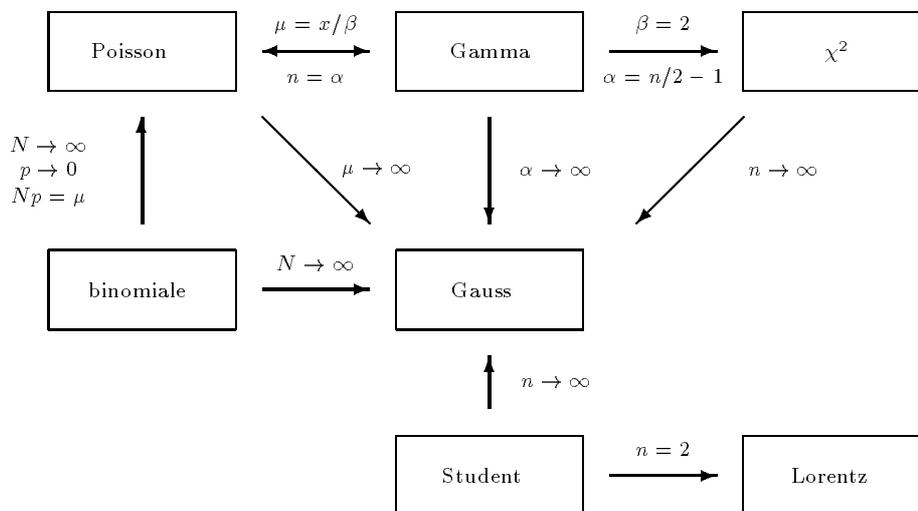


Figure 3.4 : Les relations entre les différentes distributions

une échelle de l'appareil utilisable pour la mesure de poids inconnus. Les fonctions obtenues de cette manière se présentent souvent sous la forme d'une courbe ou d'une table d'étalonnage.

Pour un appareil digital, l'incertitude de mesure est indiquée dans la description.

Pour un appareil à aiguille, la précision est caractérisée par la classe de l'appareil qui est toujours marquée sur son cadran au-dessus du symbole de position de l'appareil. L'incertitude de l'appareil est égale au produit de sa classe par la pleine échelle utilisée pour la mesure, divisé par 100 :

$$\text{incertitude} = \frac{\text{classe} \cdot \text{pleine échelle}}{100}.$$

Pour diminuer l'incertitude de mesure, il faut donc toujours travailler avec les échelles les plus sensibles possibles (les échelles qui donnent la déviation maximale acceptable). Dans la plupart des cas, on travaille avec des appareils de classe 0,5 ; 1,0 ; 1,5 ou 2,5.

Pour les expériences plus sophistiquées, cette procédure simple n'est plus suffisante. L'expérimentateur doit faire une étude approfondie du nouvel appareil pour avoir le maximum d'informations sur la fonction de résolution $S(x', x)$: vérifier si elle ne dépend que de $|x - x'|$ ou, sinon, établir la forme de cette fonction, etc.

3.4.2 ERREURS SYSTÉMATIQUES

On peut mentionner trois sources d'erreurs systématiques : la méthode de mesure choisie, le mauvais fonctionnement de l'appareillage et les erreurs d'expérimentateur. Nous allons étudier toutes ces sources d'erreurs et de voir ce qu'il faut faire dans ces cas.

Erreurs liées à la méthode de mesure

Un exemple simple d'erreur systématique provenant de la méthode de mesure est donné par la détermination d'une résistance inconnue R_x . On peut la mesurer à l'aide d'un voltmètre ayant une résistance R_V et d'un ampèremètre ayant une résistance R_A . Supposons que ces valeurs soient inconnues ; on sait seulement que R_V est grande par rapport à R_x et que R_A est petite par rapport à R_x . On branche l'ampèremètre en série avec la résistance inconnue. Le branchement du voltmètre peut être effectué de deux façons : (I) on peut mesurer la tension aux bornes de la résistance R_x (figure 3.6) ou (II) on peut mesurer la tension aux bornes de la résistance et de l'ampèremètre (figure 3.7).

Si on détermine la valeur expérimentale R_{exp} de la résistance inconnue R_x comme le rapport de la tension affichée sur le voltmètre et du courant traversant l'ampèremètre, pour ces deux branchements, on obtient les relations suivantes entre R_{exp} et R_x :

$$R_{\text{exp}}^{(I)} = R_x \frac{R_V}{R_V + R_x},$$

$$R_{\text{exp}}^{(II)} = R_x + R_A.$$

Si les appareils choisis sont de bonne qualité, pour un assez grand domaine de valeurs de la résistance R_x , telles que $R_V \gg R_x \gg R_A$, on a $R_{\text{exp}}^{(I)} \simeq R_{\text{exp}}^{(II)} \simeq R_x$. Néanmoins,

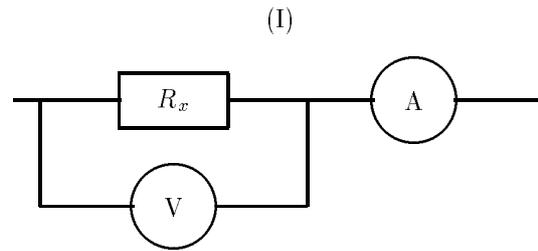


Figure 3.6 : Premier schéma possible pour mesurer la valeur d'une résistance

la première méthode donne toujours des valeurs systématiquement plus petites que la vraie valeur de R_x , tandis que la deuxième donne des valeurs systématiquement plus grandes. Dans les deux cas, on a une erreur systématique plus ou moins importante en fonction des relations entre R_V , R_A et R_x .

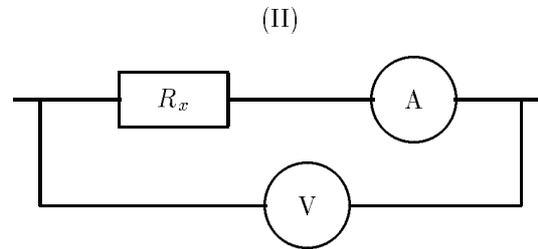


Figure 3.7 : Deuxième schéma possible pour mesurer la valeur d'une résistance

On peut donc dire que la première méthode est préférable pour mesurer des petites résistances tandis que la deuxième est plus adaptée aux grandes résistances. Cependant les deux méthodes donnent une erreur systématique qu'on ne peut éliminer qu'en connaissant les valeurs de R_V et R_A .

Proposons une troisième façon de mesurer la résistance. Pour cela, nous avons besoin d'une résistance variable dont nous pouvons établir la valeur R_v , de deux résistances identiques R et d'un appareil de mesure (d'un ampèremètre ou d'un voltmètre, au choix). Le schéma de branchement est présenté sur la figure 3.8.

Si R_x est égale à R_v , alors le courant I_a qui passe par l'ampèremètre (ou le voltmètre) est nul. On peut le voir à partir de l'expression de I_a :

$$I_a = \frac{1}{2\Delta} \frac{R_x - R_v}{R_v + R_x} I, \quad (121)$$

I étant le courant aux bornes du circuit,

$$\Delta = 1 + R_a \frac{2R + R_v + R_x}{2R(R_v + R_x)},$$

où R_a est la résistance de l'appareil (R_A ou R_V).

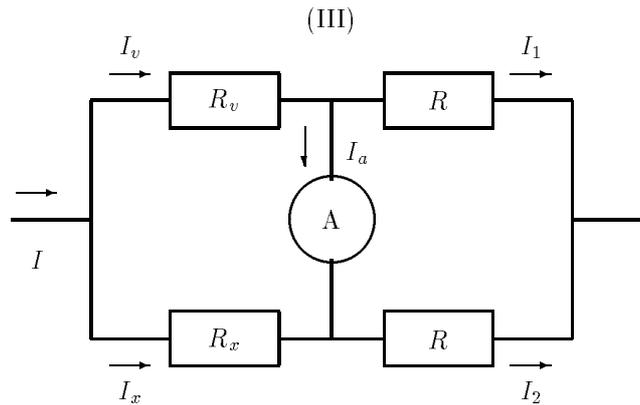


Figure 3.8 : Troisième schéma possible pour mesurer la valeur d'une résistance

L'expression (121) peut être obtenue de la façon suivante. Nous introduisons les courants I_v, I_x, I_1, I_2 (figure 3.8) et écrivons le système de 5 équations

$$\begin{cases} I = I_v + I_x \\ I = I_1 + I_2 \\ I_v = I_a + I_1 \\ I_v R_v + I_1 R = I_x R_x + I_2 R \\ I_a R_a + I_2 R = I_1 R. \end{cases}$$

Nous exprimons I_v, I_x et I_2 en fonction de I, I_a et I_1

$$\begin{cases} I_v = I_a + I_1 \\ I_x = I - I_1 - I_a \\ I_2 = I - I_1 \end{cases}$$

et obtenons deux équations

$$\begin{cases} R_v(I_a + I_1) + R I_1 = R_x(I - I_1 - I_a) - R(I - I_1) \\ R_a I_a + R(I - I_1) = R I_1. \end{cases}$$

En éliminant I_1 , il est possible d'écrire

$$I_a \frac{R_a}{2R} + I \frac{1}{2} = I \frac{R_x + R}{2R + R_x + R_v} - I_a \frac{R_x + R_v}{2R + R_x + R_v}.$$

Cette relation nous donne la formule (121).

Nous devons faire varier la résistance R_v jusqu'à annuler le courant I_a .

Quels sont les avantages d'une telle méthode par rapport aux méthodes précédentes ? Premièrement, il n'y a pas d'erreurs systématiques liées à la méthode. Si nos appareils sont précis nous obtiendrons exactement la valeur

$$R_{\text{exp}}^{(III)} = R_x.$$

Deuxièmement, nos mesures sont extrêmement simples : nous voulons annuler le courant et nous ne devons faire aucun calcul. Troisièmement, il est relativement facile de vérifier si le zéro est bien établi. Supposons que la valeur du courant est non nulle $I_a = I_0 \neq 0$, mais tellement petite que notre ampèremètre n'arrive pas à le détecter. Pour s'affranchir du problème, il suffit d'augmenter le courant extérieur I d'un facteur n , afin que le courant I_a augmente aussi d'un facteur n (voir (121)) et qu'il redevienne détectable. Ainsi nous pouvons corriger la valeur de R_v pour rétablir le zéro.

Les inconvénients possibles de cette méthode sont la difficulté de trouver une résistance variable de bonne qualité et la durée d'une telle expérience.

Dans l'exemple précédent apparaissent deux conceptions différentes d'une expérience. Dans la première approche, nous devons d'abord calibrer les appareils de mesure (voltmètre et ampèremètre) à l'aide d'étalons et ensuite les utiliser pour mesurer des valeurs physiques inconnues. Dans la deuxième approche, nous comparons directement la valeur inconnue à l'étalon. La deuxième approche est généralement plus précise mais elle est aussi plus coûteuse. Ces deux conceptions de mesure sont utilisées partout dans la vie courante. Le choix dépend de la précision recherchée et des moyens disponibles. Par exemple nous pouvons mesurer une masse, soit à l'aide d'un pèse-personne qui utilise un ressort préalablement calibré, soit à l'aide d'une balance qui équilibre la masse inconnue par des poids connus.

Erreurs liées au fonctionnement d'appareils

Le deuxième type d'erreurs systématiques est lié au mauvais fonctionnement de l'appareillage ou au changement des conditions de déroulement de l'expérience. Ces erreurs peuvent être diverses et elles dépendent de l'expérience concrète. L'exemple le plus simple est le mauvais réglage du zéro de l'appareil. Avant toute mesure il faut s'assurer que le zéro est réglé correctement. Cette vérification ne prend pas beaucoup de temps mais elle permet d'éviter des erreurs grossières et elle doit devenir une habitude pour l'expérimentateur.

L'instabilité des conditions de déroulement de l'expérience donne lieu à une dérive systématique des mesures. Par exemple la position du zéro d'un wattmètre peut varier lors d'une expérience. Un autre exemple d'une telle erreur est la mesure de la vitesse d'une boule métallique dans un liquide visqueux. Si cette expérience dure longtemps, la température du liquide peut varier avec la variation de la température ambiante et ce changement modifie la viscosité du liquide.

Erreurs d'expérimentateur

Finalement les erreurs de l'expérimentateur constituent le troisième type d'erreurs systématiques. Par exemple certaines personnes évitent tel ou tel chiffre lors des estimations de fractions de divisions d'échelle d'un appareil. Ou encore, quand on modifie les paramètres d'une expérience, le système a besoin d'un certain temps pour se mettre en équilibre et les indications des appareils peuvent être instables pendant quelques secondes. Il ne faut pas se précipiter pour faire les mesures. Lors des mesures d'un intervalle de temps, une erreur systématique peut être introduite par le fait que des personnes différentes ont des vitesses de réaction différentes.

Une erreur presque inévitable intervient lors de la lecture des indications des appareils à aiguille : il existe toujours une certaine distance entre l'aiguille et l'échelle et le résultat lu dépend de l'angle de vision. De plus, si l'aiguille se trouve entre deux divisions d'échelle, il y aura une erreur liée au choix de la valeur retenue.

Toutes ces erreurs sont presque inévitables. Il faut savoir les estimer en sachant bien que ces estimations sont personnelles, subjectives, de la responsabilité de l'expérimentateur.

3.4.3 COMMENT ÉVITER LES ERREURS SYSTÉMATIQUES ?

Pour éviter ces erreurs on peut donner quelques recommandations pratiques. Les erreurs systématiques proviennent souvent du mauvais fonctionnement de l'appareillage ou de l'expérimentateur lui-même. Ce dernier paragraphe contient quelques recommandations générales qui permettront d'éviter une grande partie de ces erreurs.

Commençons par les questions de planification et de réalisation d'une expérience sont d'une importance fondamentale. Même dans le cas d'une manipulation relativement simple en travaux pratiques il faut leur consacrer quelques minutes. Quels sont les points auxquels il faut faire attention ?

Les conditions de déroulement de l'expérience

Une manipulation dure plusieurs heures et demande un effort mental assez important. L'expérimentateur peut être fatigué et il peut se tromper. C'est pourquoi il faut commencer par la préparation de la place de travail : on ne laisse que les objets indispensables (le cahier d'expérience, la calculatrice, un stylo, etc.), l'endroit doit être bien éclairé, la température ambiante ne doit pas être trop élevée et surtout rester stable, il faut éviter les courants d'air. La stabilité de la température rend le travail plus confortable et diminue les erreurs systématiques liées aux changements des conditions de l'expérience. Il faut placer l'appareillage de façon telle que les appareils les plus fréquemment utilisés soient facilement accessibles.

Vérification des choses évidentes

Parfois, il vaut mieux vérifier des choses qui paraissent évidentes. Les appareils ne doivent pas bouger. Si la base de l'appareil est considérée comme horizontale il faut, au moins, le vérifier à l'œil nu. En optique, la condition importante est l'alignement de tous les appareils sur un même axe. Ainsi nous éviterons beaucoup d'erreurs systématiques et le processus expérimental sera accéléré. Si nous utilisons un circuit électrique alimenté directement par le réseau EDF, nous devons mesurer la tension car elle peut être différente de 220 V. Les appareils alimentés par des piles ont la "mauvaise habitude" de tomber en panne d'alimentation au moment le plus important de l'expérience. Pour éviter ce problème il faut vérifier l'état des piles avant l'expérience.

Symétrie apparente

Si le montage possède des éléments identiques, il faut les interchanger et répéter la mesure. Par exemple, sur la figure 3.8, nous avons un schéma pour déterminer une