# Mesures et incertitudes – Fiche 2 : Notion d'incertitude, d'incertitude-type et d'incertitude élargie, estimation de l'incertitude-type expérimentale par étude statistique de répétabilité (évaluation de type A)

La fiche 1 nous a permis d'identifier deux types d'erreurs qui peuvent se produire lors de la mise en place d'un protocole expérimental : les erreurs systématiques et les erreurs aléatoires.

Les erreurs systématiques sont souvent difficiles à identifier. Mais lorsque c'est le cas, il faut, dans la mesure du possible, corriger le résultat de la mesure.

En revanche les erreurs aléatoires sont toujours présentes mais vérifient des lois statistiques. Il sera donc possible de les estimer par une étude statistique de la répétabilité de la mesure. Il faudra ensuite présenter un résultat qui en tient compte.

## I. Notion d'incertitude de mesure et présentation du résultat

Du fait de la variabilité des résultats expérimentaux due aux erreurs aléatoires, le résultat final attribué à un mesurage (quelque soit la méthode de mesure) devra comporter deux valeurs numériques :

- le meilleur estimateur de la grandeur mesurée
- le meilleur estimateur de l'incertitude de mesure associé à un niveau de confiance

Pour un mesurande noté $M$, le résultat d'un mesurage devra être noté :

$$M = \widehat{m} \pm U_{p\%}(M)$$

$\widehat{m}$ : le meilleur estimateur de la grandeur mesurée

$U_{p\%}(M)$ : le meilleur estimateur de l'erreur de mesure appelée **incertitude de mesure élargie** (la notation $U$ vient de l'anglais « uncertainty ») associé à un niveau de confiance $p\%$.

**Signification :**

Cette notation signifie que la valeur vraie a $p\%$ de chance se trouver dans l'intervalle :

$$[\widehat{m} - U_{p\%}(M) \,;\, \widehat{m} + U_{p\%}(M)]$$

Cette intervalle est appelé **intervalle de confiance**.

En général, la largeur de cet intervalle est choisie pour avoir 95% ou 99% de chance de trouver la valeur vraie à l'intérieur.

Un niveau de confiance plus élevé correspondra pour un même mesurage à un intervalle plus important autour du meilleur estimateur.

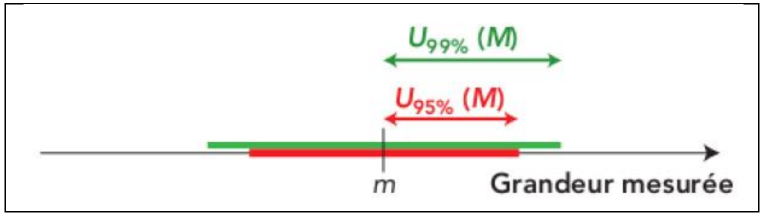



**Figure 1 : Importance du choix du niveau de confiance**

**Remarque : nombre de chiffres significatifs**

L'incertitude ne doit pas être donnée avec un nombre excessif de chiffres, à savoir : 1 ou 2 chiffres significatifs. En effet l'incertitude sur l'incertitude est assez importante : de l'ordre de 10-25 %.

En général on ne garde qu'un seul chiffre significatif en arrondissant à la valeur supérieure (on préfère majorer l'incertitude que la minorer). Mais si en ne gardant qu'un seul chiffre l'arrondi entraîne une trop grande surestimation (plus de 10% de la valeur de l'incertitude) on garde deux chiffres.

Une fois l'arrondi sur l'incertitude effectué il faut supprimer les chiffres qui n'ont pas de sens sur l'estimateur : on prendra comme dernier chiffre significatif, celui de même position que celui de l'incertitude. On arrondira l'estimateur avec les règles usuelles (cf. document de cours sur les mesures)

**Exemples :**

Mesure d'une résistance

$\hat{r} = 100{,}351389\ \Omega$ et $U_{95\%} = 0{,}842349\ \Omega$ résultat :

$$r = 100{,}4 \pm 0{,}9\ \Omega$$

Mesure d'une concentration

$\hat{r} = 0{,}1412\ \mathrm{mol.L^{-1}}$ et $U_{95\%} = 1{,}64.10^{-2}\mathrm{mol.L^{-1}}$, résultat :

$$r = 0{,}141 \pm 0{,}017\ \mathrm{mol.L^{-1}}$$

**Méthodes pour évaluer une incertitude :**

Suivant la méthode utilisée pour effectuer le calcul d'une incertitude de mesure, on peut classer cette incertitude dans l'un des deux types ci-dessous :

- une **incertitude de type A** est évaluée par des méthodes statistiques de répétabilité qui mettent en jeu la **moyenne** et l'**écart-type**. Elle est issue de l'exploitation d'un nombre important de valeurs mesurées de manière identique.
- une **incertitude de type B** est évaluée par d'autres méthodes. Elle correspond en général à une mesure unique. Sa détermination n'est pas simple, car il faut prendre en compte toutes les sources d'erreurs ou, au préalable, avoir identifié les sources d'erreurs les plus importantes.

# II. de l'incertitude expérimentale par étude statistique de répétabilité (évaluation de type A)

## 1. Valeur moyenne et écart-type expérimental des mesures

### a. Exemple

Un exemple permettra de redéfinir des grandeurs statistiques déjà vu en cours de mathématiques au lycée :

Une expérience consiste à faire croître une population de bactéries sur un support nutritif. On désire ensuite connaître la densité $D$ de ces bactéries exprimées en nombre de bactéries par millimètres carrés. Pour cela, on fait croître la population sur un support de taille 5cm × 5cm sur lequel est dessinée une grille dont chaque petit carreau fait 1mm×1mm. En utilisant un microscope on compte le nombre de bactéries par carreau sur vingt carreaux pris au hasard. On trouve les valeurs suivantes : 56, 57, 58, 58, 59, 59, 60, 60, 60, 61, 61, 61, 61, 62, 62, 62, 63, 63, 64, 65.

On peut représenter l'histogramme des valeurs à l'aide d'un tableur (LibreOffice par exemple).

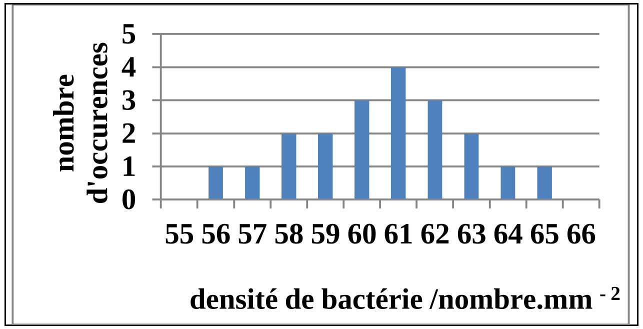



**Figure 2 : Histogramme des mesures sur Libre Office**

Pour donner toutes les informations relatives à ce mesurage il faut donner une **information sur l'estimateur le plus proche de la valeur** réelle de la densité de bactérie ainsi qu'une **information sur la dispersion des valeurs autour de cet estimateur**.

### b. Meilleur estimateur de la valeur réelle et information sur la dispersion des valeurs

**Meilleur estimateur de la grandeur mesurée $X$ :** la valeur moyenne

$$\hat{x} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$n$ : le nombre de mesures effectuées de la grandeur $X$
$\hat{x}$ : le meilleur estimateur de $x$
$x_i$ : les différentes valeurs mesurées pour la grandeur $X$
$\overline{x}$ : la valeur moyenne des $n$ mesures $x_i$

**Indicateur de la dispersion des valeurs :** l'écart-type expérimental des mesures

$$s_{\text{exp}}(X) = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}$$

$s_{\text{exp}}(X)$ : écart-type expérimental des mesures (écart-type de répétabilité) dans la même unité que $X$
$n$ : le nombre de mesure effectuée de la grandeur $X$
$x_i$ : les différentes valeurs mesurées pour la grandeur $X$
$\overline{x}$ : la valeur moyenne des $n$ mesures

Plus un écart-type est élevé plus cela signifie que les valeurs mesurées sont dispersées autour de la valeur moyenne.

Entraînez-vous à utiliser les fonctionnalités des tableurs tels que LibreOffice et de votre calculatrice pour déterminer la moyenne et l'écart-type expérimental dans le cadre de l'exemple précédent (sans avoir à entrer la formule). Voici les valeurs que vous devez obtenir :

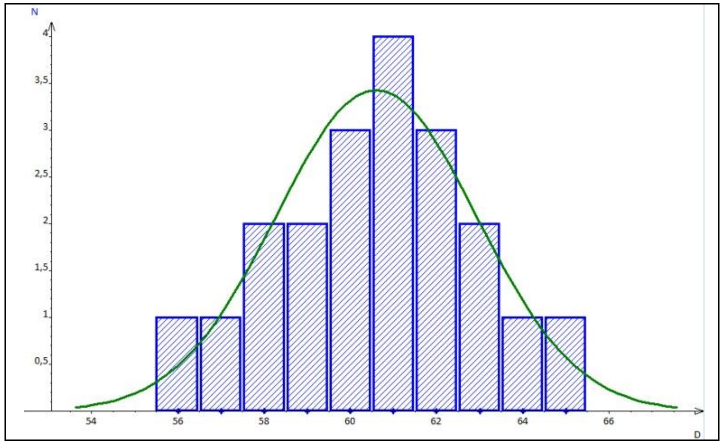
$$\overline{d} = 60{,}6 \text{ bactéries. mm}^{-2}$$

$$s_{\text{exp}}(D) = 2{,}326 \text{ bactéries. mm}^{-2}$$

### c. Loi Gaussienne

Généralement, lorsque l'on dispose d'un très grand nombre d'expériences, on constate que l'histogramme qui représente la dispersion des résultats du mesurage suit une loi gaussienne (cf. cours de statistique).

Superposons au diagramme précédent la courbe de Gauss associée :



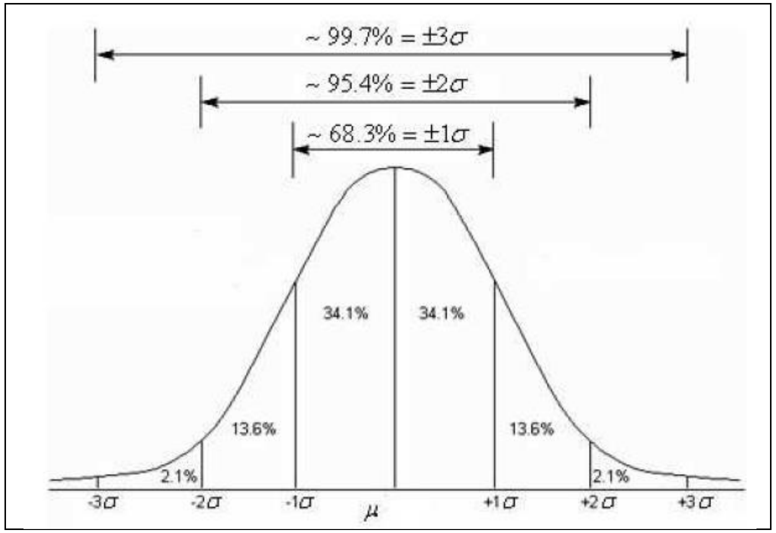

**<u>Figure 4</u> : Histogramme des mesures et courbe de Gauss associé sur Regressi**

Une répartition gaussienne est symétrique par rapport à sa valeur la plus probable appelé espérance $\mu$ et les faibles écarts à cette valeur sont plus probables que les forts, elle suit l'expression suivante (qui n'est pas à retenir) :

$$p(x) = \frac{n}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

L'écart-type permet de définir que la probabilité qu'une mesure $x$ de $X$ se situe dans l'intervalle $[\mu - \sigma\,;\mu + \sigma]$ est de 68,3%.

Dans l'exemple des bactéries $x$ représente l'ensemble les valeurs que peut prendre la grandeur $D$, l'espérance $\mu$ est la valeur moyenne $\overline{d}$. $\sigma$ est l'écart-type de la gaussienne qui permet en quelque sorte de définir la largeur de la courbe de Gauss.




**<u>Figure 5</u> : Espérance et écart-type pour une gaussienne**

**Remarque :**

Dans l'exemple de la densité des bactéries, la répartition des mesures expérimentales ne correspond pas parfaitement à la répartition gaussienne attendue. En effet la répartition gaussienne ne peut être obtenue que si l'on effectue une infinité de mesures.

Comme nous ne pouvons jamais effectuer une infinité de mesure, l'écart-type expérimental des mesures $s_{\exp}(x)$ ne sera qu'un estimateur de l'écart-type réel $\sigma(x)$ de la distribution gaussienne des valeurs.

## 2. Ecart-type sur la valeur moyenne : incertitude-type

L'écart-type calculé précédemment donne une estimation moyenne de la distance que sépare une valeur mesurée quelconque de la valeur moyenne. Comment à partir de ce paramètre en déduire une incertitude sur la valeur moyenne obtenue ?

Il semble évident que le nombre de mesures effectuées va jouer un rôle sur l'incertitude de mesure sur la moyenne (c'est-à-dire l'estimateur de l'erreur de mesure entre la valeur moyenne et la valeur vraie du mesurande). Il est donc communément admis que :

Plus un mesurage comporte un grand nombre de mesure, plus la valeur moyenne obtenue est précise

Comparons ces deux diagrammes simulés à l'aide d'un programme informatique : on effectue un certain nombre de mesures d'un mesurande dont la valeur vraie vaut 100, ces mesures sont soumises à une erreur aléatoire imposée par le logiciel. Dans le premier cas, seule 5 mesures sont effectuées, dans le deuxième, 30.

En jaune : la valeur moyenne
En rouge : la valeur vraie

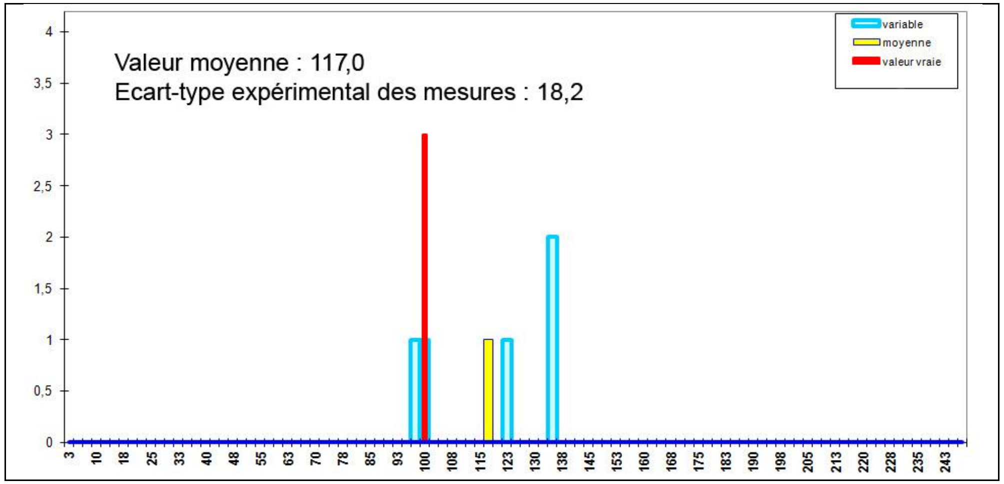



**Figure 7: Simulation de 5 mesures expérimentales**

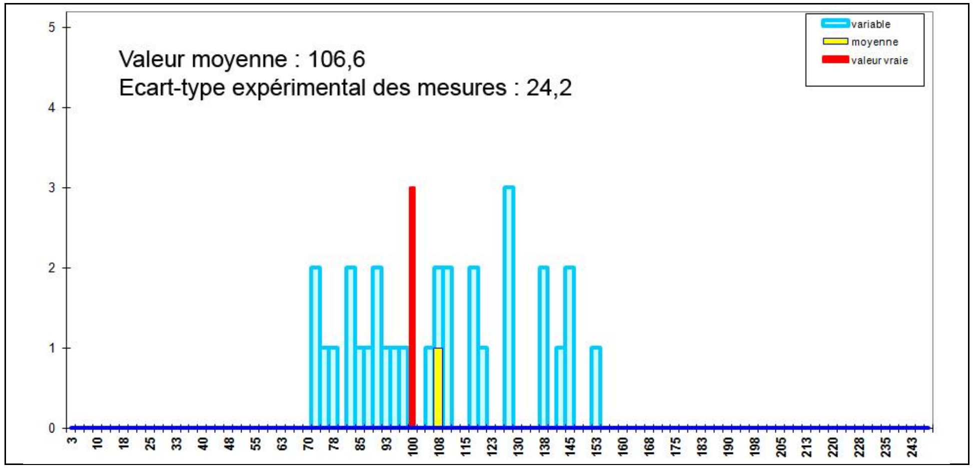



**Figure 8 : Simulation de 30 mesures expérimentales**

La deuxième simulation permet d'obtenir une moyenne plus précise (plus proche de la valeur vraie) pourtant l'écart-type expérimental des mesures est plus important que celui de la première simulation.

La grandeur qui va permettre d'obtenir une estimation de l'incertitude sur la moyenne n'est pas l'écart-type des mesures mais l'**écart-type des moyennes**.

Imaginons que nous effectuons 200 fois les mêmes opérations de mesurage comportant $n$ mesures ($n = 5$ ou $30$), superposons toutes les mesures et pour chaque échantillon de $n$ mesures, notons la moyenne.

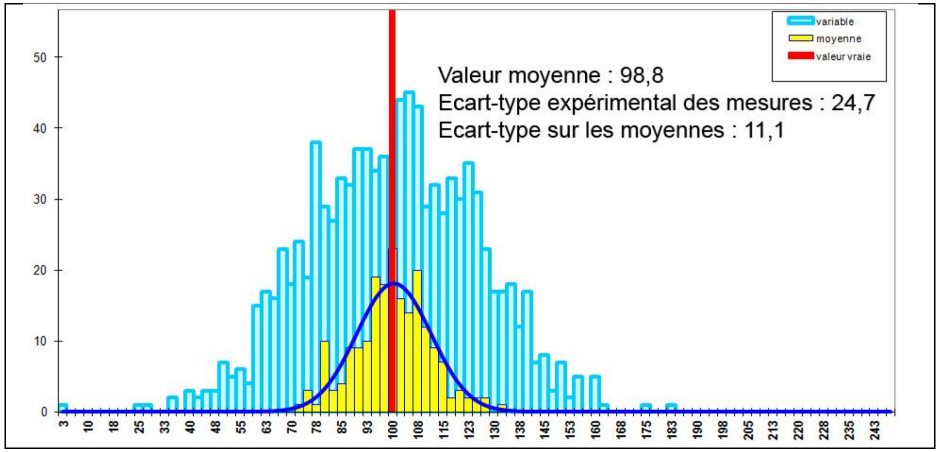



**Figure 9 : Simulation 200 échantillons de 5 mesures**

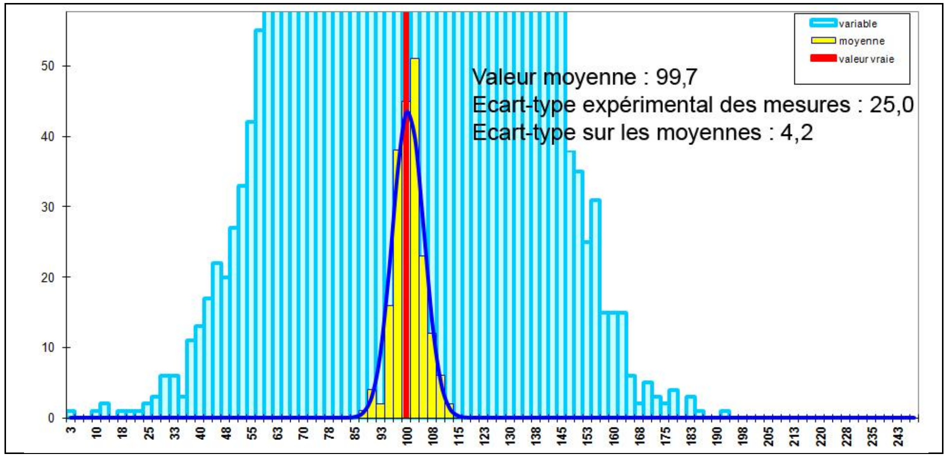



**Figure 10 : Simulation 200 échantillons de 30 mesures**

Considérons maintenant la dispersion des valeurs moyennes de chaque échantillon : les échantillons de 5 mesures sont plus dispersés que ceux de 30 mesures : leur écart-type est plus important. Mais dans les deux cas la répartition des valeurs moyennes respecte une allure gaussienne.

**Définition :**

La grandeur qui va permettre d'obtenir une estimation de l'incertitude sur la valeur moyenne d'un échantillon est l'écart-type des moyennes, appelée appelée **incertitude-type** et notée $\boldsymbol{u(X)}$.

Elle se calcule grâce à la formule suivante (qui peut se démontrer) :

$$u(X) = \frac{s_{\exp}(X)}{\sqrt{n}}$$

$s_{\exp}(X)$ : l'écart-type expérimental des mesures de l'échantillon de $n$ mesures
$n$ : nombre de mesures effectuées

Remarque : la formule donne bien une incertitude-type sur la moyenne d'autant plus faible que le nombre de mesures permettant d'obtenir cette moyenne est important.

### 3. Incertitude élargie et facteur d'élargissement

Nous venons de définir une grandeur caractéristique de la dispersion des valeurs moyennes, mais il faut définir une incertitude liée à un intervalle de confiance.

#### a. Dans le cas d'une répartition parfaitement gaussienne

| **Niveau de confiance $\boldsymbol{p\%}$** | 68,3 | 95,4 | 99,7 |
|---|---|---|---|
| **Incertitude élargie : $\boldsymbol{U_{p\%}(X)}$** | $u(x)$ | $2u(x)$ | $3u(x)$ |

Si un mesurande suit une loi de répartition gaussienne de moyenne $\overline{x}$ et d'incertitude-type sur la moyenne $u(X)$, on peut montrer que la probabilité d'obtenir un résultat compris dans l'intervalle $\left[\overline{x} - u(X) \,;\, \overline{x} - u(X)\right]$, si on réalise à nouveau une série de $n$ mesures pour laquelle on calcule la moyenne, est de 68,3%. Pour que l'intervalle de confiance ait plus de signification, on préfère prendre un niveau de confiance plus élevé.

On parle alors d'**incertitude élargie** et le coefficient devant $u(x)$ est appelé **facteur d'élargissement**.

#### b. Dans le cas réel

En réalité, comme on effectue rarement une infinité de mesures, on ne dispose que d'estimateurs $\overline{x}$ et $u(x)$, on introduit alors des facteurs d'élargissement corrigés (facteurs de Student) $t_n^{p\,\%}$ en fonction du nombre de mesures effectuées. Ex. à 95 % :

| $n$ | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 14 | 16 | 18 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $t_n^{95\,\%}$ | 3,18 | 2,78 | 2,57 | 2,45 | 2,37 | 2,31 | 2,26 | 2,20 | 2,16 | 2,13 | 2,11 | 2,09 |

#### c. Comment présenter le résultat

Pour un niveau de confiance à 95%, on a :

$$X = \overline{x} \pm t_n^{95\,\%} u(X) = \overline{x} \pm t_n^{95\,\%} \frac{s_{\exp}(X)}{\sqrt{n}}$$

Si $n$ est très grand on fera l'approximation gaussienne :

$$X = \overline{x} \pm 2\frac{s_{\exp}(X)}{\sqrt{n}}$$

**Application à l'exemple de la densité de bactérie :** la densité de bactérie, pour un indice de confiance à 95 % en faisant l'hypothèse d'une répartition gaussienne, vaut :

$$D = (60{,}6 \pm 1{,}1)\ \text{bactéries.mm}^{-2}$$

$$\text{car}\quad U_{95\%}(D) = 2 \times \left(s_{\exp}(X)/\sqrt{20}\right) = 1{,}040\ \text{bactéries.mm}^{-2}$$