Alternatives afin d'estimer le résultats d'un petit nombre de mesures d'une grandeur

1. Problématique, hypothèses

Soit un nombre restrinct d'échantillons ...

Soit un petit échantillon de mesure d'une même grandeur, contenant entre 2 et 20 points, comme il se produit souvent dans des cas concrets de mesurage en laboratoire, issu d'une **population normale** d'espérance $\mu := 10$ et d'écart-type $\sigma := 1.5$.(Vous pouvez changer les valeurs de ces paramètres) :

$N := 7$: taille de l'échantillon $\quad$ $\text{normal} := \text{rnorm}(N, \mu, \sigma)$

$$\text{normal}^T = (9.342 \quad 8.981 \quad 9.29 \quad 8.573 \quad 7.471 \quad 10.065 \quad 9.819)$$

Nous allons rechercher les meilleurs estimateurs calculables sur l'échantillon pour obtenir une estimation convenable de la "moyenne" (positionnement) et de "l'écart" (incertitude) de la mesure.

"écart-type"

Afin de comparer plusieurs cas possibles, nous allons considérer, en plus du cas normal donné auparavant, 3 autres échantillons qui diffèrent peu d'un échantillon normal, mais dont la différence peut être significative, selon la méthode d'analyse utilisée. L'intérêt réside précisément dans la comparaison des estimateurs appliqués aux échantillons.

1.1 Échantillon comportant une valeur aberrante

Ce cas est le plus souvent rencontré. Bien qu'il existe plusieurs tests statistiques pour déterminer si une valeur est aberrante ou non, il est toujours très difficile de décider de retirer cette valeur, surtout si la mesure s'est déroulée très soigneusement, et si la valeur aberrante est relativement proche des autres points. Nous allons donc considérer l'effet d'une valeur aberrante (très marquée, pour l'exemple) sur la dernière valeur mesurée, qui est "accidentellement" multipliée par 10:

$$\text{aberrant} := \text{normal}$$

$$\text{aberrant}_{\text{last(aberrant)}} := 10 \cdot \text{normal}_{\text{last(normal)}}$$

$$\text{aberrant}^T = (9.342 \quad 8.981 \quad 9.29 \quad 8.573 \quad 7.471 \quad 10.065 \quad 98.19)$$

1.2 Échantillon issu d'une loi log-normale

entre

Ce cas décrit concrètement une mesure plus souvent qu'il n'y paraît. Il est celui qui rentre en ligne de compte lors d'une loi exponentielle ou logarithmique (dB), mais aussi quand l'incertitude relative (le coefficient de variation), soit le rapport σ/μ est constant. Cela revient à dire que ce cas décrit les mesures pour lesquelles l'erreur en % est constante. Pour simplifier, nous gardons la moyenne égale à μ en introduisant son logarithme:

$$\text{lognormal} := \text{rlnorm}(N, \ln(\mu), \ln(\sigma))$$

$$\text{lognormal}^T = (12.531 \quad 24.319 \quad 13.881 \quad 14.91 \quad 14.185 \quad 14.495 \quad 13.137)$$

1.3 Échantillon comportant un mélange de deux populations normales

Cette situation est réelle lorsque l'on dispose de deux échantillons de tailles différentes, pris dans dans des conditions de mesure différentes. Par exemple, un premier échantillon est pris sur un petit nombre de mesures, afin d'estimer le résultat, dans un premier temps. Puis un second échantillon est mesuré, cette fois avec plus de points, de manière plus complète. Supposons pour fixer les idées, que le premier échantillon comporte 10% des mesures, et que son écart-type (incertitude) soit 10 fois plus important que celui du second échantillon. Cela se produit notamment lorsque l'appareil de mesure est éloigné de sa résolution, l'accumulation de points pouvant diminuer l'écart sur la moyenne. Dans ce cas, on peut décrire la situation par l'approximation:

$$\text{melange} := 90\%\cdot\text{rnorm}(N, \mu, \sigma) + 10\%\cdot\text{rnorm}(N, \mu, 10\cdot\sigma)$$

$$\text{melange}^T = (7.815 \quad 10.93 \quad 8.613 \quad 11.075 \quad 11.647 \quad 8.072 \quad 9.027)$$

2. Estimation de l'espérance

Dans la plupart des cas, le résultat final de la mesure est donné par la moyenne arithmétiqu (éventuellement corrigée d'une erreur systématique, que nous n'envisageons pas ici) des points de l'échantillon. C'est aussi la manière de procéder de toutes les normes. Il faut cependant se méfier de cette pratique quasi automatique, car la moyenne arithmétique est très sensible aux points exceptionnels de l'échantillon. On dit que la moyenne n'est pas un estimateur robuste de la position (l'espérance) de l'échantillon. Nous allons donc comparer plusieurs solutions. Fixons d'abord le nombre de chiffres significatifs utiles à cette comparaison:

$$\text{arrondi} := 10^{\text{floor}\left[\log\left[\sqrt{\frac{1}{2\cdot(N-1)}}\cdot\frac{\sigma}{\sqrt{N}}\right]\right]} \qquad \text{arrondi} = 0.1$$

2.1 Moyenne arithmétique

La moyenne arithmétique de l'échantillon est la méthode standard (préconisée dans les normes internationales) d'évaluation de l'espérance (du résultat final). Sous Mathcad, elle correspond à la fonction *mean* (en anglais). Construisons le tableau des moyennes:

$\text{TablM}_{0,1} := \text{"normal"}$ $\text{TablM}_{0,2} := \text{"aberrant"}$

$\text{TablM}_{0,3} := \text{"lognormal"}$ $\text{TablM}_{0,4} := \text{"mӑlange"}$

$\text{TablM}_{1,0} := \text{"moyenne"}$

$\text{TablM}_{1,1} := \text{Round}(\text{mean}(\text{normal}), \text{arrondi})$

$\text{TablM}_{1,2} := \text{Round}(\text{mean}(\text{aberrant}), \text{arrondi})$

$\text{TablM}_{1,3} := \text{Round}(\text{mean}(\text{lognormal}), \text{arrondi})$

$\text{TablM}_{1,4} := \text{Round}(\text{mean}(\text{melange}), \text{arrondi})$

$$\text{TablM} = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mӑlange"} \\ \text{"moyenne"} & 9.1 & 21.7 & 15.4 & 9.6 \end{pmatrix}$$

On remarque que la moyenne arithmétique est très sensible à la présence d'une valeur aberrante, notamment. Elle l'est aussi dans le cas de la loi log-normale, dans une moindre mesure. (**Vous pouvez en tout temps simuler d'autres échantillons en tapant simultanément Ctrl+F9**).

2.2 Médiane

L'utilisation de la médiane de l'échantillon, (qui est la valeur centrale des points de l'échantillon ordonnés en ordre croissant, si N est impair, et la moyenne des deux points les plus proches de la valeur centrale si N est pair), est plus rare. Rappelons cependant que pour toute loi de probabilité symétrique, la moyenne et la médiane sont égales. Sous Mathcad, la médiane est donnée par la fonction *median*:

$TablM_{2,0} := \text{"mйdiane"}$

$TablM_{2,1} := \mathrm{Round}(\mathrm{median}(normal), arrondi)$

$TablM_{2,2} := \mathrm{Round}(\mathrm{median}(aberrant), arrondi)$

$TablM_{2,3} := \mathrm{Round}(\mathrm{median}(lognormal), arrondi)$

$TablM_{2,4} := \mathrm{Round}(\mathrm{median}(melange), arrondi)$

$$TablM = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"moyenne"} & 9.1 & 21.7 & 15.4 & 9.6 \\ \text{"mйdiane"} & 9.3 & 9.3 & 14.2 & 9 \end{pmatrix}$$

Contrairement à la moyenne, la médiane est moins sensible aux valeurs aberrantes. Il s'agit donc d'une alternative intéressante de calcul du résultat final, du moins pour des échantillor dont la taille est très petite.

2.3 Par l'utilisation de l'étendue

Cette estimation de la position du résultat à partir de la connaissance de l'étendue est plus rare. En fait, il ne s'agit pas à proprement parler de l'étendue, qui est la différence entre la valeur max et min de l'échantillon, mais bien que la valeur centrale basée sur max et min. Elle est très sensible aux valeurs exceptionnelles. Sous Mathcad, il n'existe pas de fonction donnant directement l'estimation de l'espérance à partir de l'étendue de l'échantillon. Il est très simple de la créer:

mais plutôt de la valeur ...

$$Rmean(v) := \frac{\max(v) + \min(v)}{2}$$

Et de l'appliquer aux échantillons:

$TablM_{3,0} := \text{"max/min"}$

$TablM_{3,1} := \mathrm{Round}(Rmean(normal), arrondi)$

$$\text{TablM}_{3,2} := \text{Round}(\text{Rmean}(\text{aberrant}), \text{arrondi})$$

$$\text{TablM}_{3,3} := \text{Round}(\text{Rmean}(\text{lognormal}), \text{arrondi})$$

$$\text{TablM}_{3,4} := \text{Round}(\text{Rmean}(\text{melange}), \text{arrondi})$$

$$\text{TablM} = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"moyenne"} & 9.1 & 21.7 & 15.4 & 9.6 \\ \text{"mйdiane"} & 9.3 & 9.3 & 14.2 & 9 \\ \text{"max/min"} & 8.8 & 52.8 & 18.4 & 9.7 \end{pmatrix}$$

On remarque que son application n'est pas recommandable (sauf pour N=2), les résultats variant beaucoup par rapport au cas normal, notamment la valeur aberrante.

3. Estimation de l'incertitude du résultat

Un résultat de mesure ne saurait avoir d'intérêt sans la donnée de son incertitude. L'incertitude est (selon la norme EN) la valeur caractérisant la dispersion des valeurs pouvant raisonnablement être attribuées à la mesure. Nous admettons ici que l'écart-type statistique est une bonne mesure de l'incertitude du résultat. Il n'en est pas toujours ainsi, notamment lorsqu'interviennent la résolution de la mesure, ou sa dérive au cours du temps.
L'incertitude est le plus souvent estimée sur la base de l'écart-type de l'échantillon. Nous analysons ici d'autres variantes, qui permettent une comparaison de sensibilité, notamment pour les petits échantillons.
Nous nous limitons à la détermination de l'incertitude de l'espérance de l'échantillon, c'est-à-dire l'incertitude de la valeur de position, à ne pas confondre avec l'écart-type de la loi elle-même.

3.1 Par l'écart-type expérimental

L'écart-type expérimental est l'écart-type basé sur le calcul de l'échantillon, (avec une pondération N-1 pour rattraper son biais, d'où la majuscule dans sa définition *Stdev*), qui estime l'écart-type de la loi, soit $\sigma = 1.5$:

$$\text{Stdev}(\text{normal}) = 0.864 \qquad \text{Stdev}(\text{aberrant}) = 33.738$$

$$\text{Stdev}(\text{lognormal}) = 4.036 \qquad \text{Stdev}(\text{melange}) = 1.579$$

$$\text{TablU} := \text{submatrix}(\text{TablM}, 0, 0, 0, 4)$$

On remarque tout de suite que cet écart est assez éloigné de sa valeur attendue σ, dans presqu tous les cas. Il s'agit d'un effet du à la taille faible de l'échantillon.

L'incertitude sur le résultat de mesure est lié à cet écart. En fait, il est connu que la loi de probabilité de la moyenne s'approche (pour de grands échantillons) d'une loi normale (même si la loi originale n'est pas normale) d'écart-type donné par:

$$u_{minf}(v) := \frac{\mathrm{Stdev}(v)}{\sqrt{\text{длина}(v)}} \qquad \mathrm{Round}\left(\frac{\sigma}{\sqrt{N}}, \mathrm{arrondi}\right) = 0.6$$

donc numériquement, pour les cas particuliers, on obtient alors:

$$\mathrm{TablU}_{1,0} := \text{"ET:racine"}$$

$$\mathrm{TablU}_{1,1} := \mathrm{Round}\left(u_{minf}(\mathrm{normal}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{1,2} := \mathrm{Round}\left(u_{minf}(\mathrm{aberrant}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{1,3} := \mathrm{Round}\left(u_{minf}(\mathrm{lognormal}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{1,4} := \mathrm{Round}\left(u_{minf}(\mathrm{melange}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU} = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"ET:racine"} & 0.3 & 12.8 & 1.5 & 0.6 \end{pmatrix}$$

On s'aperçoit que l'estimation de l'incertitude selon cette méthode semble sous estimée dans le cas normal, et très dépendante du type de l'échantillon. La raison est à rechercher dans le fait que, bien que la variance expérimentale soit un estimateur non-biaisé de la variance de la population, l'écart-type expérimental est un estimateur biaisé de l'écart-type de la population. En introduisant la correction suivante, qui compense ce biais:

$$u_{mc}(v) := \left| \begin{array}{l} n \leftarrow \text{длина}(v) \\ u \leftarrow \dfrac{\text{Stdev}(v)}{\sqrt{n}} \cdot \dfrac{\Gamma\left[\dfrac{(n-1)}{2}\right] \cdot \sqrt{n-1}}{\Gamma\left[\dfrac{(n)}{2}\right] \cdot \sqrt{2}} \end{array} \right.$$

On obtient:

$$TablU_{2,0} := \text{"ET:corrigй"}$$

$$TablU_{2,1} := \text{Round}\left(u_{mc}(\text{normal}), \text{arrondi}\right)$$

$$TablU_{2,2} := \text{Round}\left(u_{mc}(\text{aberrant}), \text{arrondi}\right)$$

$$TablU_{2,3} := \text{Round}\left(u_{mc}(\text{lognormal}), \text{arrondi}\right)$$

$$TablU_{2,4} := \text{Round}\left(u_{mc}(\text{melange}), \text{arrondi}\right)$$

$$TablU = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"ET:racine"} & 0.3 & 12.8 & 1.5 & 0.6 \\ \text{"ET:corrigй"} & 0.3 & 13.3 & 1.6 & 0.6 \end{pmatrix}$$

L'estimation de l'incertitude est légèrement plus grande que dans le cas précédent.
Pour obtenir une estimation convenable de l'incertitude sur la moyenne, il est bien connu que la loi t de Student en donne une approximation correcte si la loi est normale.
On peut montrer que, pour un risque de $\alpha := 15.87\%$, on a une incertitude (correspondant à 1 écart-type de la loi normale), de:

$$u_T(v) := u_{minf}(v) \cdot qt(1 - \alpha, \text{длина}(v) - 1)$$

Donc :

$$TablU_{3,0} := \text{"ET:Student"}$$

$$TablU_{3,1} := \text{Round}\left(u_T(\text{normal}), \text{arrondi}\right)$$

$$TablU_{3,2} := Round\left(u_T(aberrant), arrondi\right)$$

$$TablU_{3,3} := Round\left(u_T(lognormal), arrondi\right)$$

$$TablU_{3,4} := Round\left(u_T(melange), arrondi\right)$$

$$TablU = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"ET:racine"} & 0.3 & 12.8 & 1.5 & 0.6 \\ \text{"ET:corrigй"} & 0.3 & 13.3 & 1.6 & 0.6 \\ \text{"ET:Student"} & 0.4 & 13.9 & 1.7 & 0.7 \end{pmatrix}$$

La correction est faible dès que la taille de l'échantillon N>6. Les 3 méthodes basées sur l'écart-type expérimental ne sont pas très différentes dans leurs résultats, en tenant compte que, pour le nombre limités de points, leurs coefficients de variation valent environ $\sqrt{\frac{1}{2\cdot(N-1)}} = 29\%$. Elles ne conviennent pas dans le cas de la présence de valeurs aberrantes, notamment.

3.2 Estimation de l'incertitude par la médiane

Pour l'estimation de l'incertitude, on utilise le fait que la relation approximative (asymptotique) suivante existe entre la variance de la médiane et la variance théorique, dans le cas de la loi normale:

$$var(med) = \frac{\pi}{2\cdot N}\cdot var(X)$$

En définissant la médiane de l'écart absolu (mea) par l'expression (On calcule ici la variance des écarts par rapport à la variance, sans tenir compte de leurs signes):

$$mea(x) := median\left(\overrightarrow{\left|x - median(x)\right|}\right)$$

On montre facilement que cette définition est reliée à l'écart-type σ de la loi normale par l'expression:

$$mea = qnorm\left(\frac{3}{4}, 0, 1\right)\cdot\sigma$$

Finalement , l'incertitude (l'écart-type) sur la médiane est donnée par la relation:

$$u_{med}(x) = \sqrt{\frac{\pi}{2 \cdot N}} \cdot \frac{mea(x)}{qnorm\left(\frac{3}{4}, 0, 1\right)} \qquad c_{med} := \sqrt{\frac{\pi}{2}} \cdot \frac{1}{qnorm\left(\frac{3}{4}, 0, 1\right)}$$

qui devient:

$$u_{med}(x) := c_{med} \cdot \frac{mea(x)}{\sqrt{\text{длина}(x)}} \qquad c_{med} = 1.86$$

L'application aux échantillons donne:

$$TablU_{4,0} := \text{"Med:mea"}$$

$$TablU_{4,1} := Round\left(u_{med}(normal), arrondi\right)$$

$$TablU_{4,2} := Round\left(u_{med}(aberrant), arrondi\right)$$

$$TablU_{4,3} := Round\left(u_{med}(lognormal), arrondi\right)$$

$$TablU_{4,4} := Round\left(u_{med}(melange), arrondi\right)$$

$$TablU = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"ET:racine"} & 0.3 & 12.8 & 1.5 & 0.6 \\ \text{"ET:corrigй"} & 0.3 & 13.3 & 1.6 & 0.6 \\ \text{"ET:Student"} & 0.4 & 13.9 & 1.7 & 0.7 \\ \text{"Med:mea"} & 0.4 & 0.5 & 0.5 & 0.9 \end{pmatrix}$$

Cette méthode a l'avantage important de corriger la valeur aberrante (pour N>2) et souvent les écarts de la loi log-normale. Elle reste applicable pour les autres cas, avec plus ou moins de succès.

Il existe une autre approche relativement simple du calcul de l'incertitude basée sur la médiane. Cette méthode n'est applicable que pour des échantillons pour lesquels N>4. Elle fait appel à la loi binomiale. En effet, si B est une variable binomiale de probabilité de succ $p := 0.5$, alors, pour tout entier k entre 0 et $\text{floor}\left(\frac{N}{2}\right)$, on peut facilement calculer la probabilité que B se situe entre k et N-k. Par analogie avec la loi normale, cette probabilité doit être égale à $\beta := 50\% - \frac{68.27\%}{2}$, soit un intervalle de 2 σ. Il faut donc rechercher la valeur entière k qui s'approche le plus de la probabilité de 68.27% par excès:

$k := \text{qbinom}(\beta, N, p) - 1 \cdot (N > 7)$ $\qquad k = 2$

$\gamma_M := \text{pbinom}(N - k, N, p) - \text{pbinom}(k, N, p)$ $\qquad \gamma_M = 71\%$

$\gamma_m := \text{pbinom}[N - (k + 1), N, p] - \text{pbinom}(k + 1, N, p)$ $\qquad \gamma_m = 27\%$

Comme ces valeurs sont relativement éloignées de l'intervalle voulu, il est nécessaire de corriger le tir par une sorte d'extrapolation, par le calcul de:

$$I := \frac{\gamma_M - 68.27\%}{\gamma_M - \gamma_m} \qquad \lambda := \frac{(N - k) \cdot I}{k + (N - 2 \cdot k) \cdot I}$$

Les limites de l'intervalle d'incertitude sont calculées par:

```
u_med2(v) := | x ← sort(v)
             | u ← 0  if k < 1
             | otherwise
             |    | haut ← λ·x_(N−1−k) + (1 − λ)·x_(N−1−k+1)
             |    | bas ← λ·x_(k+1) + (1 − λ)·x_k
             |    | u ← haut − bas
             | u
```

et :

$$\mathrm{TablU}_{5,0} := \text{"Med:binome"}$$

$$\mathrm{TablU}_{5,1} := \mathrm{Round}\left(u_{med2}(\mathrm{normal}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{5,2} := \mathrm{Round}\left(u_{med2}(\mathrm{aberrant}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{5,3} := \mathrm{Round}\left(u_{med2}(\mathrm{lognormal}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU}_{5,4} := \mathrm{Round}\left(u_{med2}(\mathrm{melange}), \mathrm{arrondi}\right)$$

$$\mathrm{TablU} = \begin{pmatrix} 0 & \text{"normal"} & \text{"aberrant"} & \text{"lognormal"} & \text{"mйlange"} \\ \text{"ET:racine"} & 0.3 & 12.8 & 1.5 & 0.6 \\ \text{"ET:corrigй"} & 0.3 & 13.3 & 1.6 & 0.6 \\ \text{"ET:Student"} & 0.4 & 13.9 & 1.7 & 0.7 \\ \text{"Med:mea"} & 0.4 & 0.5 & 0.5 & 0.9 \\ \text{"Med:binome"} & 0.7 & 0.9 & 0.9 & 2.4 \end{pmatrix}$$

Bien qu'applicable indépendamment de toute hypothèse sur la loi de l'échantillon, cette méthode, relativement difficile à mettre en œuvre sans ordinateur, donne des résultats certainement trop larges pour la plupart des cas pratiques.

3.3 Estimation de l'incertitude par l'étendue

L'utilisation de l'étendue, définie par la relation:

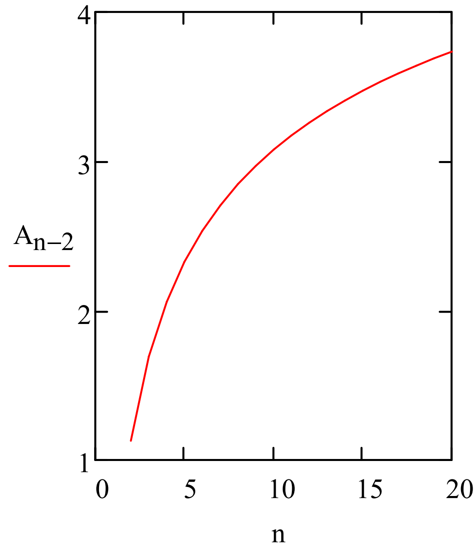
$$\mathrm{range}(x) := \max(x) - \min(x)$$

est admise dans la plupart des normes, afin de simplifier le calcul de l'incertitude. Ces mêm normes donnent aussi des tables de coefficients A, dépendants de la taille N de l'échantillor afin de déterminer l'écart-type à partir de l'étendue:

$A :=$

	0
0	1.128
1	1.693
2	2.059
3	2.326
4	2.534
5	2.704
6	...

$n := 2 .. 20$



$$u_e(x) := \frac{\text{range}(x)}{\sqrt{\text{длина}(x)} \cdot A_{\text{длина}(x)-2}}$$

$$TablU_{6,0} := \text{"Etendue"}$$

$$TablU_{6,1} := \text{Round}\left(u_e(\text{normal}), \text{arrondi}\right)$$

$$TablU_{6,2} := \text{Round}\left(u_e(\text{aberrant}), \text{arrondi}\right)$$

$$TablU_{6,3} := \text{Round}\left(u_e(\text{lognormal}), \text{arrondi}\right)$$

$$TablU_{6,4} := \text{Round}\left(u_e(\text{melange}), \text{arrondi}\right)$$

TablU =

0	"normal"	"aberrant"	"lognormal"	"mйlange"
"ET:racine"	0.3	12.8	1.5	0.6
"ET:corrigй"	0.3	13.3	1.6	0.6
"ET:Student"	0.4	13.9	1.7	0.7
"Med:mea"	0.4	0.5	0.5	0.9
"Med:binome"	0.7	0.9	0.9	2.4
"Etendue"	0.4	12.7	1.6	0.5

Cette méthode donne des résultats comparables à celles utilisant l'écart-type expérimental. Elle a l'avantage d'être très rapide d'emploi, du moins pour les échantillons de petites tailles Elle est évidemment très sensible aux valeurs aberrantes.

4. Conclusions

1. Pour de très petits échantillons (N<4), la moyenne arithmétique n'est pas forcément le meilleur estimateur de l'espérance, la médiane est comparable. Dans ce cas, l'écart-type de la moyenne doit être évalué à l'aide de la loi de t, ou par calcul avec la médiane, plus rapide. Le calcul à l'aide de l'étendue est également possible, si l'échantillon ne contient pas de valeurs aberrantes ou suspectes.
2. Pour des tailles d'échantillons plus importantes, mais toujours modestes, la moyenne arithmétique devient préférable pour l'estimation de la position, avec l'accroissement de la taille N; dans ce cas, l'estimation de l'écart-type est donnée par l'écart-type expérimental sur $\sqrt{N}$, ou par la méthode de l'étendue.
3. Dans tous les cas, l'utilisation de deux méthodes différentes pour estimer la valeur centrale et l'incertitude est recommandée. La méthode standard (moyenne + écart-type expérimental / $\sqrt{N}$) est à utiliser en premier. Ensuite, la méthode basée sur la médiane (médiane + mea/$\sqrt{N}$) est comparée à la précédente. Une différence de l'ordre de 25% entre les deux méthodes doit alerter l'expérimentateur sur l'une des anomalies suivantes:
   - L'échantillon s'écarte fortement d'un échantillon normal;
   - Il y a probablement des valeurs aberrantes;
   - Les valeurs centrales et/ou incertitudes n'ont pas de sens dans le cas étudié.

5. Références

- Konstantin Protassov, *Analyse statistique des données expérimentales*, EDP Sciences, Les Ulis, 2002.
- Norme NF ENV 13005, *Guide pour l'expression de l'incertitude de mesure*, Afnor, 1999.
- Rand R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, San Diego, 1997
- Jörg W. Müller, *Possible Advantages of a Robust Evaluation of Comparisons*, in J. Res. Natl. Inst. Stand. Technol. 105, pp. 551-555, (2000)

6. Commentaires et remarques

Vos commentaires sur cette feuille de comparaison sont les bienvenus à l'adresse suivante:
paratte @ ieee.org.

$\nu := 0 .. N - 1$