



CERD



The BiotechNet event / Événement BiotechNet

Kempinski Palace, Djibouti

25 Octobre 2022

MODULE 1 - Statistiques basiques

Pr. Tarik Ainane

OBJECTIFS :

- 1 - Formation théorique des notions fondamentales autour de la biostatistique descriptive.
- 2 - Tests statistiques
- 3 - Exemples en métrologie.

STATISTIQUE DESCRIPTIVE

INDICATEURS DE POSITION

Moyenne

Valeur maximale

Valeur minimale

Mode

Médiane

Quartile Q1

Quartile Q2



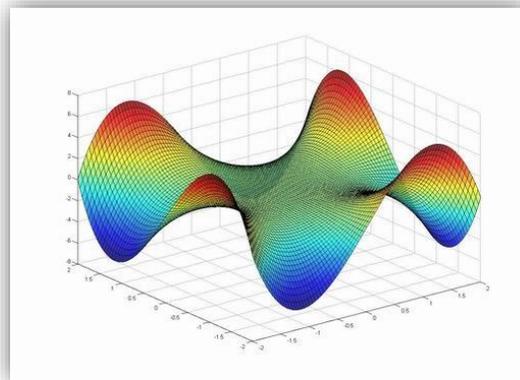
INDICATEUR DE DISPERSION

L'écart-type

Le coefficient de variation

L'écart interquartile

L'étendue



INDICATEURS DE POSITION :

Soit une mesurande M de n valeurs mesurées : x_1, x_2, \dots, x_n .

La moyenne :

La **moyenne arithmétique** est la somme des valeurs de la variable divisée par le nombre d'individus :

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

➤ x_i : Mesures individuelles.

➤ n : Nombre de mesures.

La valeur maximale :

La **valeur maximale** x_{max} est la plus grande valeur prise par les mesures.

La valeur minimale :

La **valeur minimale** x_{min} est la plus petite valeur prise par les mesures.

La médiane :

La **médiane** est la valeur de la série qui partage la distribution en deux sous-ensembles d'égale effectifs c'est à dire qu'une moitié de la série est inférieure à la médiane et l'autre moitié est supérieure à la médiane.

INDICATEURS DE DISPERSION :

L'écart-type :

L'écart-type noté σ est la mesure de dispersion la plus couramment utilisée en statistique lorsqu'on emploie la moyenne pour calculer une tendance centrale. Il mesure donc la dispersion autour de la moyenne :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Coefficient de variation :

L'écart-type observé est un paramètre de dispersion absolue qui mesure la variation absolue des mesures indépendamment de l'ordre de grandeur des données. *Le coefficient de variation* noté « C.V » est un indice de dispersion relatif est égal à :

$$C.V = \frac{\sigma}{\bar{x}} \times 100 \quad (\%)$$

L'écart interquartile :

L'écart interquartile ou *l'étendue interquartile* « EI » est une mesure de dispersion qui s'obtient en faisant la différence entre le troisième et le premier quartile :

$$E.I = Q_3 - Q_1$$

L'étendue R :

L'étendue « R » est une mesure de dispersion qui s'obtient en faisant la différence entre la *valeur maximum* et la *valeur minimum* :

$$R = x_{max} - x_{min}$$

ACTIVITÉ :



Mesurer une distance?

En sciences expérimentales il n'existe pas de mesures exactes. Les erreurs de mesure peuvent provenir de plusieurs sources :

- ❖ qualité des instruments ;
- ❖ l'expérimentateur (l'erreur est humaine) ;
- ❖ la variabilité de la grandeur mesurée (par exemple une masse d'un solvant chimique se change au cours du temps selon leur condition d'évaporation) ;
- ❖ ...

Par conséquent, lors du mesurage d'une grandeur, on évaluera son incertitude.

Méthode de mesure

Succession logique des opérations, décrites d'une manière générique, mises en œuvre lors de l'exécution de mesurages.

Mesurage

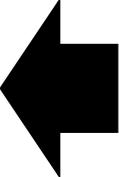
Ensemble d'opérations ayant pour but de déterminer la valeur d'une grandeur.

Résultat de mesurage

Valeur attribuée à un mesurande, obtenue par mesurage.

Valeur

Expression quantitative d'une grandeur particulière, généralement sous la forme d'une unité de mesure multipliée par un nombre

 **Valeur Vraie?**

Valeur vraie (d'une grandeur)

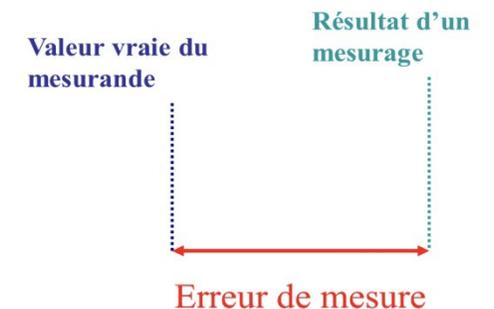
Valeur compatible avec la définition d'une grandeur particulière donnée. C'est une valeur que l'on obtiendrait par un mesurage parfait. Toute valeur vraie est par nature indéterminée.

Erreur de mesure

Résultat d'un mesurage moins valeur vraie du mesurande.

$$e_i = x_i - x_0$$

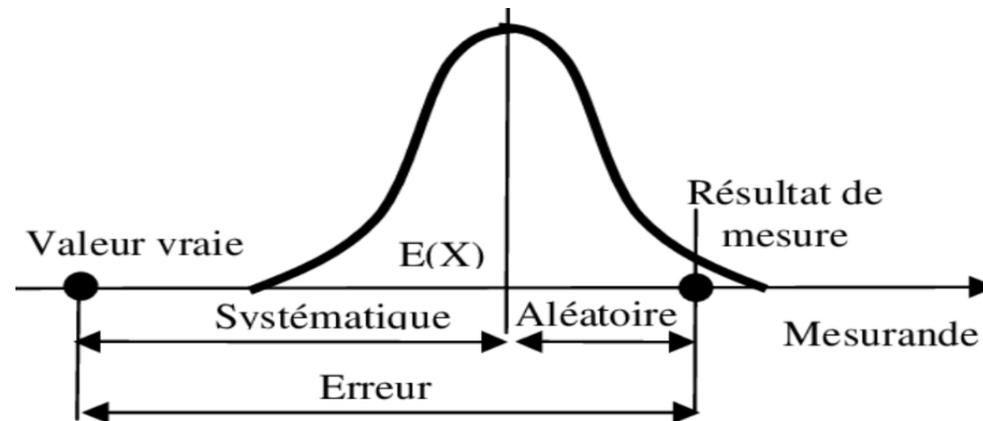
Avec x_i est le résultat de mesurage et x_0 est la valeur vraie de la grandeur.



Deux composantes d'erreurs de mesure :

- ✚ Erreur systématique e_s : erreur constante, si les facteurs d'influence sont constants.
- ✚ Erreurs aléatoires e_a

Erreur de mesure = erreur aléatoire + erreur systématique



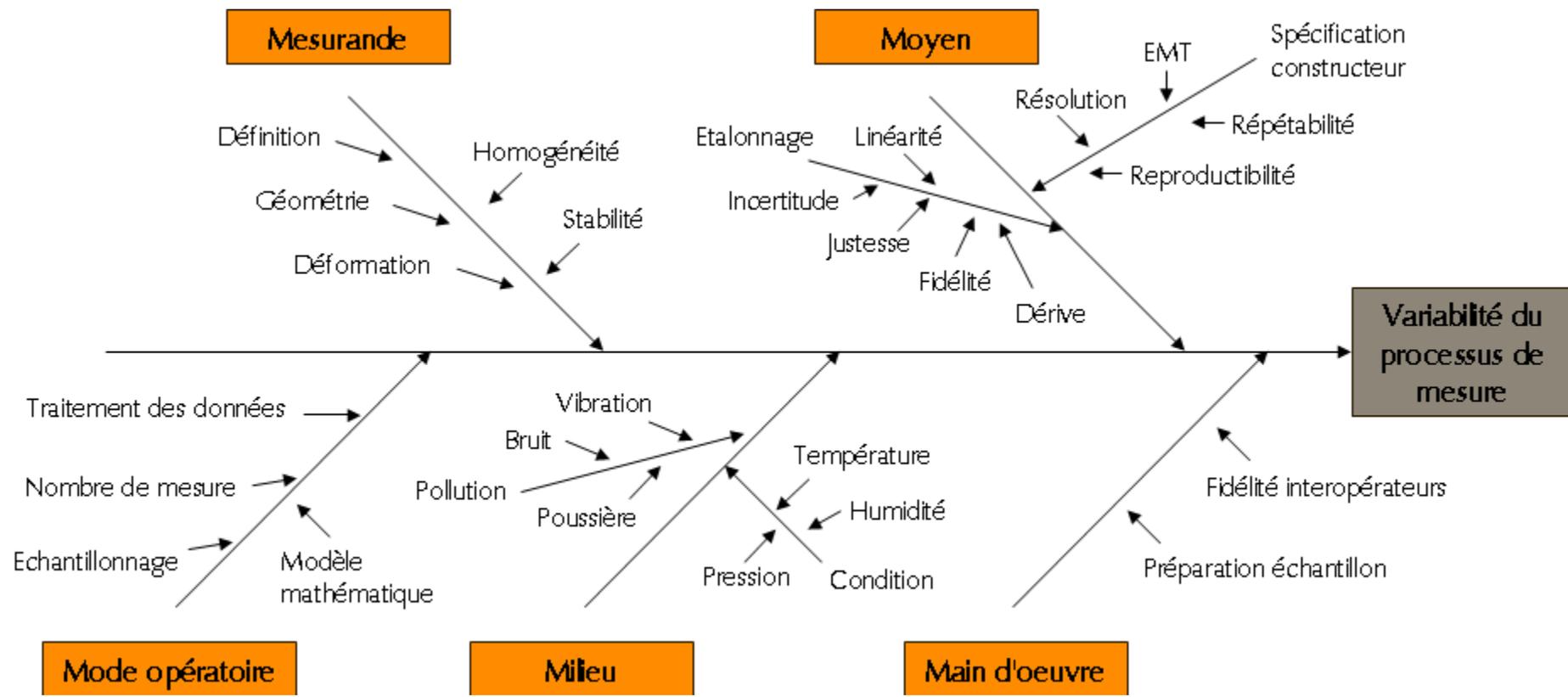
Erreur systématique

Moyenne qui résulterait d'un nombre infini de mesurages du même mesurande, effectués dans les conditions de répétabilité, moins une valeur vraie du mesurande. $e_s = \mu - x_0$

Les erreurs systématiques sont dues aux conditions de répétabilité et restent constantes en fonction du temps.

L'erreur systématique caractérise la justesse de la série de mesure.

Erreur de justesse d'un instrument de mesure = erreur systématique



Erreur aléatoire

Résultat d'un mesurage moins la moyenne d'un nombre infini de mesurages du même mesurande, effectués dans les conditions de répétabilité.

$$e_a = x_i - \mu$$

L'incertitude de mesure a été définie comme un paramètre qui caractérise la dispersion des valeurs autour d'une valeur "moyenne" d'un mesurande.

Résultat du mesurage = (valeur numérique \pm **incertitude) unité**

Selon le mode d'obtention on distingue :

Evaluation de Type A

Incertitude de type A

Evaluation de Type B

Incertitude de type B

EVALUATION DE TYPE A :

Méthode d'évaluation de l'incertitude par analyse statistique d'une série d'observations.

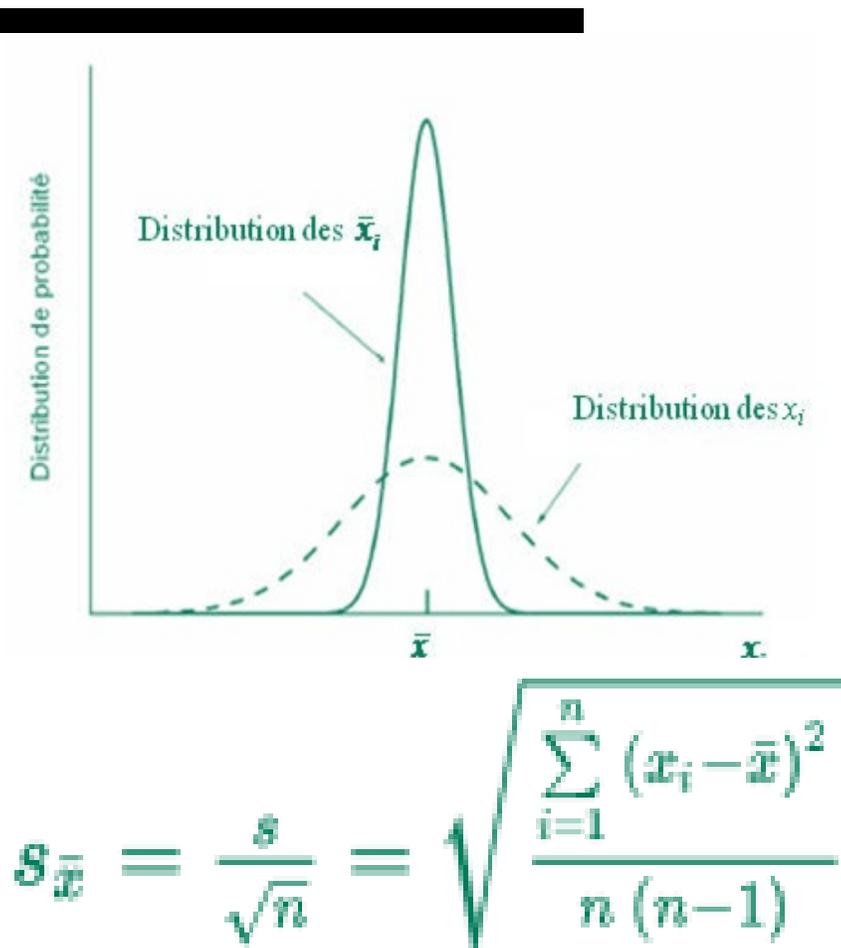
Une série de n observation de valeur x_i avec un moyenne \bar{X} la dispersion autour de cette moyenne permet de définir **l'incertitude de type A** .

En utilisant des grandeurs statistique comme la **demi-étendue** or **l'écart-type** .

La valeur déterminé de définit que la valeur mesuré x situer dans l'intervalle suivant :

$\bar{X} \pm S_{\bar{x}}$ avec un intervalle de confiance 66%

$\bar{X} \pm 2S_{\bar{x}}$ avec un intervalle de confiance de 95%



EVALUATION DE TYPE B :

Méthode d'évaluation de l'incertitude par des moyens autres que l'analyse statistique de séries d'observations, en générale en se base sur des données fournis par le constructeur.

Incertitudes associées aux principales erreurs de manipulation

Erreur identifiée	Incertitude associée
Information sur une solution étalonnée	0,1%
Volume extrait à partir d'un matériel jaugé (pipettes, fioles ou ballons, ...)	0,2% du volume
Volume extrait à partir d'un matériel gradué (pipettes, burettes, ...)	0,5% plus petite graduation
Lecture sur un cadran numérique	Unité du dernier chiffre significatif stable

INCERTITUDE DE TYPE A :

L'**incertitude de type A** est évaluée lorsque :

- ❖ Un même manipulateur réalise plusieurs fois le mesurage d'un *même mesurande M* dans les mêmes conditions expérimentales.
- ❖ ou plusieurs manipulateurs réalisent simultanément *la même mesure* avec du matériel similaire (mesure de pH d'une solution par plusieurs binômes de TP par exemple).

L'écart type permet d'évaluer **l'incertitude-type $u(M)$** . Pour « n » mesures **l'incertitude-type $u(M)$** vaut :

$$u(M) = \frac{\sigma_{n-1}}{\sqrt{n}}$$

L'incertitude élargie appelée **incertitude de répétabilité $U(M)$** est égale au produit de **l'incertitude-type $u(M)$** par un facteur k appelé **facteur d'élargissement** :

$$U(M) = k \cdot u(M) = k \cdot \frac{\sigma_{n-1}}{\sqrt{n}}$$

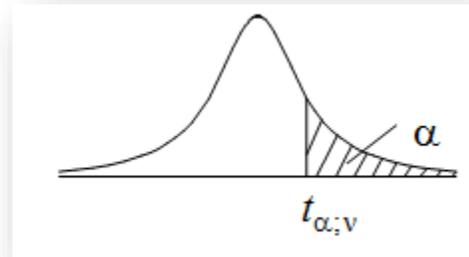
Le résultat **du mesurage** s'écrit :

$$M = \bar{x} \pm U(M)$$

Le **coefficient d'élargissement k** dépend de :

- du niveau de confiance α (par exemple 95% ou 99%) ;
- du nombre n de mesures.

Le **coefficient d'élargissement k** est donné par la **loi de Student**.
 La détermination de k fait par l'intersection entre la ligne de degré de liberté $\nu = n-1$ et la colonne de probabilité P ($P=1-2\alpha$).



ν	Niveau de confiance α				
	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
∞	1.282	1.645	1.960	2.326	2.576

Exemple :

on effectue $n = 10$ mesures du pH d'une eau de rivière, l'écart-type expérimentale vaut $0,15$, et la moyenne des mesures est $6,51$.

Pour un niveau de confiance de 95%, quel est le résultat du mesurage ainsi que l'intervalle de confiance ?

❖ L'incertitude de répétabilité $U(M)$ vaut :

$$U(M) = k \cdot u(M) = k \cdot \frac{\sigma_{n-1}}{\sqrt{n}} = 2,26 \cdot \frac{0,15}{\sqrt{10}} = 0,11$$

❖ Résultat du mesurage :

$$pH = \bar{x} \pm U(M) = 6,51 \pm 0,11$$

La valeur vraie du pH de l'eau à une probabilité de 95% de se trouver dans l'intervalle de confiance suivant :

$$pH \in [6,40;6,62]$$

INCERTITUDE DE TYPE B :

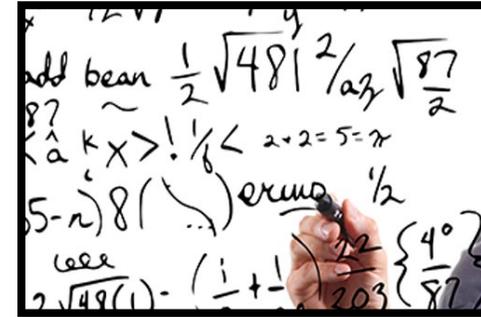
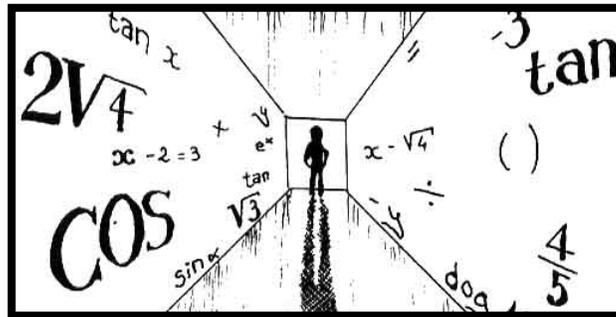
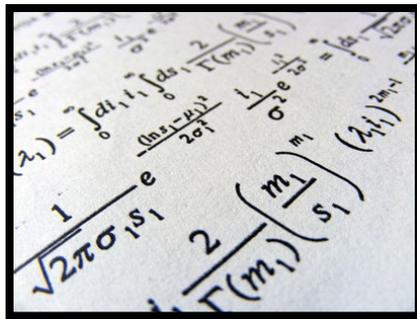
L'évaluation d'une **incertitude de type B** s'effectue lors d'une mesure unique. On cherchera à évaluer l'incertitude du à la précision de chaque l'appareil comme :

- ❖ la résolution (graduation ou digit) qui correspond à l'incertitude de lecture ;
- ❖ la tolérance du constructeur ;
- ❖ l'incertitude de l'étalon ;
- ❖ les grandeurs ayant une influence sur la mesure (température, hygrométrie...).

Pour les **incertitudes de type B**, on utilise l'incertitude élargie :

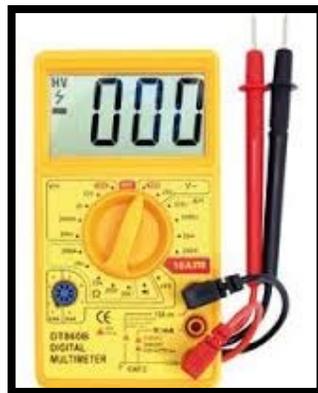
$$U = k \cdot u$$

Pour un niveau de confiance de **95 %**, le facteur d'élargissement est **k = 2** et pour **99 %** on a **k = 3**.





Incertitude liée à la tolérance t d'un appareil



Incertitude liée à la précision d'un appareil à affichage numérique



Incertitude de lecture

Incertitude de type B

Incertitude de lecture :

Lors d'une mesure par lecture sur une échelle ou un cadran, ***l'incertitude-type*** de lecture est donnée par la formule :

$$u_{\text{lecture}} = \frac{1 \text{ graduation}}{\sqrt{12}}$$

Exemple :

Un *thermomètre* est gradué en degré Celsius (*1 graduation correspond à 1 degré Celsius*).

• *Evaluer l'incertitude élargie de lecture U pour un niveau de confiance de 99% ?*

$$U = k \cdot u_{\text{lecture}} = k \cdot \frac{1 \text{ graduation}}{\sqrt{12}}$$

$$U = 3 \times \frac{1}{\sqrt{12}} = 0,87^\circ\text{C}$$



Incertitude liée à la tolérance t d'un instrument :

L'incertitude-type est donnée par la relation :

$$u_{\text{tolérance}} = \frac{t}{\sqrt{3}}$$

Exemple :

Une fiole jaugée de 50 ml à une tolérance $t = \pm 0,1$ ml.

• Calculer l'incertitude élargie U liée à la tolérance de la fiole avec un niveau de confiance de 95% ?

$$U = k \cdot u_{\text{tolérance}} = k \cdot \frac{t}{\sqrt{3}}$$

$$U = 2 \times \frac{0,1}{\sqrt{3}} = 0,12 \text{ ml}$$



Incertitude liée à la précision d'un instrument à affichage numérique :

La notice de *l'instrument* indique généralement *la précision p* correspondant à un pourcentage de la valeur lue sur l'écran et par un certain nombre de digit (plus petite valeur possible affichée).

L'incertitude-type liée à la précision de cet appareil est :

$$u_{\text{précision}} = \frac{p}{\sqrt{3}}$$

Exemple :

Une balance à une précision de 0,01 g. Il affiche la valeur 2,34 g.

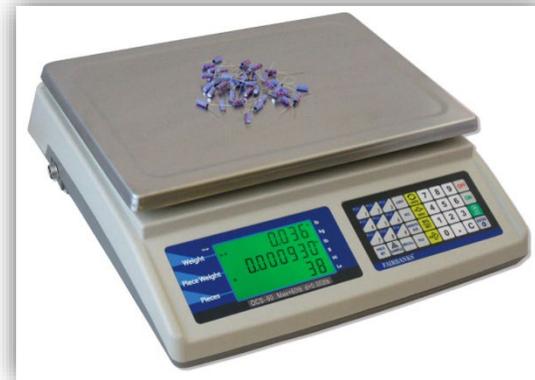
• Calculer l'incertitude élargie relative à la précision de l'appareil correspondant à un niveau de confiance de 99%. Quel est le résultat du mesurage ?

$$U = k \cdot u_{\text{précision}} = k \cdot \frac{p}{\sqrt{3}}$$

$$U = 3 \times \left(\frac{0,01}{\sqrt{3}} \right) = 0,02 \text{ g}$$

• Résultat du mesurage :

$$m = 2,34 \text{ g} \pm 0,02 \text{ g}.$$



Propagation de la dispersion :

Si on a des **valeurs x_i** d'une **mesure** ont été obtenues par plusieurs **dispersions U_i** (erreurs indépendants), les incertitudes se combinent aléatoirement de sorte que **l'incertitude sur la mesure combinée U** est donné par :

$$U = \sqrt{\sum_i U_i^2}$$

Exemple :

Calcul de l'incertitude type B d'étalonnage d'un instrument de pesage : Balance à indication numérique :

- ❖ Incertitude liée à la Résolution : $U_r = 0,001 \text{ g}$
- ❖ Incertitude liée à l'influence de la Température : $U_t = 0,006 \text{ g}$
- ❖ Incertitude liée à la grandeur appliquée : $U_e = 0,002 \text{ g}$
- ❖ Incertitude liée à l'Effet de la masse volumique de l'air : $U_{pa} = 0,001 \text{ g}$
- ❖ Incertitude liée à l'Effet de l'excentration de charge (corps) : $U_{ex} = 0,003 \text{ g}$

Incertitude total vaut :

$$U = \sqrt{U_r^2 + U_t^2 + U_e^2 + U_{pa}^2 + U_{ex}^2}$$

$$U = 0,007 \text{ g}$$

FIN