

Service: Qualité des laboratoires

Traitement des valeurs censurées

Date de mise à jour : 30/09/2019

TABLE DE MATIERES

1	INTRODUCTION	3
2	POSITION DU PROBLEME	3
3	NOUVELLE METHODE	4
4	EXEMPLES	6
5	RAPPORT GLOBAL	9
6	RAPPORT INDIVIDUEL	9
7	RAPPORT INDIVIDUEL ANNUEL	10
8	COMPARAISON DES DEUX METHODES	11
9	PROBLEME DES PETITS GROUPES	15
10	RECOMMANDATIONS FINALES	15

1 INTRODUCTION

Il arrive que la concentration d'un analyte dans un spécimen contrôle EEQ ne puisse être mesurée car elle tombe sous le seuil de sensibilité de la méthode analytique utilisée. On rend alors un résultat nul ou inférieur à la limite de sensibilité, par exemple <0.10 . On dit que le résultat est censuré à gauche (Left-censored observation). Généralement, les valeurs censurées sont systématiquement éliminées de l'échantillon et le traitement statistique, c'est-à-dire le calcul des percentiles P_{25} , P_{50} (médiane) et P_{75} , s'effectue sur l'ensemble des données non censurées. Cette façon de procéder n'est pas optimale car elle entraîne une perte non négligeable d'information. Une nouvelle méthode d'évaluation des résultats EEQ est proposée qui tient compte des valeurs censurées (à gauche et aussi à droite), lorsque de telles valeurs sont présentes dans l'échantillon des résultats fournis par les laboratoires.

2 POSITION DU PROBLEME

Considérons l'ensemble des résultats EEQ fournis par n laboratoires pour un constituant X. Cet échantillon peut contenir des données censurées (à gauche).

Le problème statistique consiste à estimer les quartiles P_{25} , P_{50} et P_{75} à partir des données, afin d'obtenir des estimations "robustes" de la valeur "vraie" (M) du constituant X dans le spécimen contrôle et de la variabilité interlaboratoire (SD) des résultats.

On peut alors écrire :

$$M = P_{50}$$
$$SD = 0.74 \times (P_{75} - P_{25})$$

Notons que si P_{25} n'existe pas, alors $SD = 1.48 \times (P_{75} - P_{50})$. Connaissant M et SD, on peut calculer la distance entre le résultat (R) obtenu par un laboratoire et la valeur centrale (M). En réalité, on calcule le Z-score, c'est-à-dire

$$Z = (R - M)/SD$$

On peut aussi calculer le U-score (%) qui exprime l'écart relatif entre le résultat (R) et la valeur centrale (M). On a

$$U = (R - M)/M \quad (\times 100\%)$$

Le programme informatique EEQ utilise, depuis sa mise en application en 1989, des formules connues pour estimer les quartiles. Celles-ci cependant ne tiennent pas compte des valeurs censurées. La méthode présentée dans ce rapport permet d'estimer les quartiles en présence de valeurs censurées à gauche. En outre, elle fournit les mêmes résultats que l'ancienne méthode lorsqu'il n'y a pas de valeurs censurées dans l'échantillon.

3 NOUVELLE METHODE

La méthode proposée consiste à estimer la fonction de répartition (cumulative distribution function) de X et à calculer ensuite les quartiles à partir de cette fonction de répartition. Soit $\{x_1, \dots, x_n\}$ les résultats EEQ fournis par les n laboratoires. On procède en six étapes.

Etape 1: Trier l'échantillon par ordre croissant. Si une valeur censurée (à gauche) est égale à une valeur non censurée, elle doit la précéder dans la séquence ascendante. Par exemple <0.2 doit être classé avant 0.2 .

Etape 2: On répertorie dans l'échantillon ainsi trié, toutes les valeurs non censurées distinctes. Désignons par $X_{(0)} < X_{(1)} < \dots < X_{(k-1)} < X_{(k)}$ les k valeurs distinctes non censurées ainsi obtenues où, par convention, $X_{(0)} = -\infty$. Notons que k peut varier entre 0 (lorsque toutes les valeurs de l'échantillon sont censurées – cas extrême) et n (lorsqu'il n'y a aucune valeur censurée dans l'échantillon et que toutes les valeurs sont distinctes).

On note N_i le nombre d'observations de l'échantillon d'effectif n inférieures ou égales à $X_{(i)}$ et $R_{(i)}$ le nombre d'observations non censurées égales à $X_{(i)}$, pour $i = 1, \dots, k$.

Etape 3: La fonction de répartition $F(x)$ de la variable X dans l'échantillon d'effectif n est obtenue par la formule de récurrence (descendante):

$$F(X_{(k)}) = 1$$
$$F(X_{(k-i)}) = F(X_{(k-i+1)}) \times \frac{N_{(k-i+1)} - R_{(k-i+1)}}{N_{(k-i+1)}} \quad (i = 1, \dots, k)$$

Il s'agit d'une fonction étagée qui comporte k escaliers. En effet, elle change de valeur en chaque $X_{(i)}$, ($i = 1, \dots, k$). On peut reporter cette fonction de répartition sur un graphique (voir exemples ci-après).

Etape 4: On calcule le percentile P_{75} . A cet effet, on recherche à partir de $X_{(k)}$ les deux valeurs $X_{(j-1)}$ et $X_{(j)}$ dont les valeurs de la fonction de répartition $F(X_{(j-1)})$ et $F(X_{(j)})$ sont telles que $F(X_{(j-1)}) \leq 0.75 < F(X_{(j)})$. Deux cas peuvent se présenter:

- si $F(X_{(j-1)}) = 0.75$, alors $P_{75} = (X_{(j-1)} + X_{(j)})/2$
- sinon, $P_{75} = X_{(j)}$

Il peut arriver que P_{75} n'existe pas. Dans ce cas, il en est de même pour P_{25} et P_{50} et SD n'est pas estimable non plus.

Remarque:

Lorsque l'échantillon ne contient pas de valeurs censurées et que l'effectif n est un multiple de 4 moins 1 ($n = 4 \times m - 1$), alors on utilise systématiquement la formule $P_{75} = X_{(j)}$, sauf si $F(X_{(j-1)}) = (3n - 1)/4n$, dans ce cas on calcule $P_{75} = (X_{(j-1)} + X_{(j)})/2$.

Etape 5: On calcule ensuite la médiane $M = P_{50}$ de la même manière, en recherchant les valeurs $X_{(j-1)}$ et $X_{(j)}$ telles que $F(X_{(j-1)}) \leq 0.5 < F(X_{(j)})$. Si P_{50} n'existe pas, il en est de même de P_{25} et SD n'est pas calculable.

Etape 6: On calcule ensuite le percentile P_{25} de la même manière en recherchant les valeurs $X_{(j-1)}$ et $X_{(j)}$ telles que $F(X_{(j-1)}) \leq 0.25 < F(X_{(j)})$. Si P_{25} n'existe pas, on peut estimer SD par la formule $SD = 1.48 \times (P_{75} - P_{50})$

Remarque:

Lorsque l'échantillon ne contient pas de valeurs censurées et que l'effectif n est un multiple de 4 moins 1 ($n = 4 \times m - 1$), alors on utilise systématiquement la formule $P_{25} = X_{(j)}$, sauf si $F(X_{(j)}) = (n + 1)/4n$, auquel cas on calcule $P_{75} = (X_{(j)} + X_{(j+1)})/2$.
Pour rappel, $X_{(j)}$ est tel que $F(X_{(j-1)}) \leq 0.25 < F(X_{(j)})$.

Remarques

- Ainsi présentée, la nouvelle méthode est applicable en toutes circonstances et conduit aux mêmes résultats (en tenant compte de la note pour les percentiles P_{25} et P_{75}) que l'ancienne méthode lorsqu'il n'y a pas de valeurs censurées.
- Les valeurs censurées (à gauche) supérieures à $X_{(k)}$ ne sont pas prises en considération.
- S'il existe une valeur censurée (à gauche) inférieure à $X_{(1)}$, alors la fonction de répartition n'est pas égale à 0 mais à moins l'infini (voir exemple 2 ci-après).
- L'approche développée peut s'accoutumer de la présence dans l'échantillon de valeurs censurées à droite (par exemple, >16). Lors du tri des données (étape 1), si une valeur censurée à droite est égale à une valeur non censurée, elle doit la suivre dans la séquence croissante et non la précéder! On procède ensuite comme précédemment.

4 EXEMPLES

Illustrons la méthode par quelques exemples.

Exemple 1

Considérons d'abord les résultats de digoxine en nmol/L (échantillon n°1 de l'enquête EEQ 1998/4, méthode n°37) fournis par $n=70$ laboratoires.

Données brutes

0.4	0.3	0.3*	0.0	0.3	0.3	0.4	0.3	0.3*	0.4
0.2	0.1	0.1	0.4	0.2	0.3	0.3*	0.3	0.0	0.3
0.4*	0.4*	0.4*	0.4	0.1	0.4*	0.4*	0.3	0.0	0.3*
0.3	0.4*	0.4*	0.1	0.4*	0.0	0.3	0.4	0.3	0.5*
0.0	0.3	0.2	0.2	0.3	0.3	0.5	0.2	0.0	0.6*
0.0	0.4*	0.3*	0.0	0.3*	0.4	0.3	0.2	0.4	0.2
0.3	0.4	0.3	0.2	0.3	0.5	0.4*	0.1*	0.1	0.2

Les données censurées (à gauche) sont suivies d'un astérisque.

Données triées par ordre croissant

0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	<0.1	0.1	0.1	0.1	0.1	0.1	0.2
0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	<0.3	<0.3	<0.3	<0.3	<0.3	<0.3	0.3
0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
0.3	0.3	<0.4	<0.4	<0.4	<0.4	<0.4	<0.4	<0.4	<0.4	<0.4	<0.4	0.4	0.4	0.4
0.4	0.4	0.4	0.4	0.4	0.4	<0.5	0.5	0.5	<0.6					

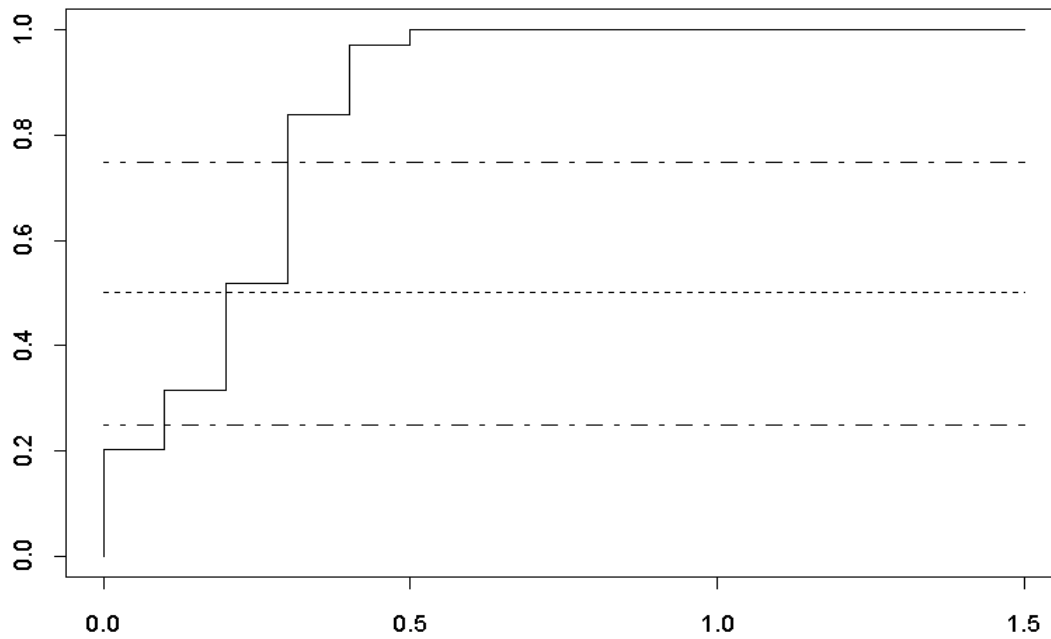
On constate que l'échantillon possède 19 valeurs censurées et 51 valeurs non censurées. On répertorie également $k=6$ valeurs distinctes non censurées, respectivement 0.0, 0.1, 0.2, 0.3, 0.4 et 0.5.

Fonction de répartition

L'application de la méthode décrite à la section 3 (étapes 2 et 3) conduit à la fonction de répartition suivante.

i	$X_{(i)}$	N_i	R_i	$F(X_{(i)})$
0	$-\infty$	0	0	0.00
1	0.0	8	8	0.20
2	0.1	14	5	0.31
3	0.2	23	9	0.52
4	0.3	47	18	0.84
5	0.4	66	9	0.97
6	0.5	69	2	1.00

Graphiquement, la fonction de répartition est illustrée sur la figure ci-dessous.



Percentiles

L'application des étapes (4-6) de la méthode conduit aux quartiles suivants:

$$P_{75} = 0.3$$

$$P_{50} = 0.2$$

$$P_{25} = 0.1$$

On a donc $M = 0.2 \text{ nmol/L}$

$$SD = 0.74 \times (0.3 - 0.1) = 0.148 \text{ nmol/L}$$

Exemple 2

Le deuxième exemple, généré artificiellement, montre que, lorsque l'échantillon ne contient pas de valeurs censurées, on retrouve les mêmes résultats qu'avec l'ancienne méthode.

Données brutes

21 17 18 4 1 7 27 13 12 23

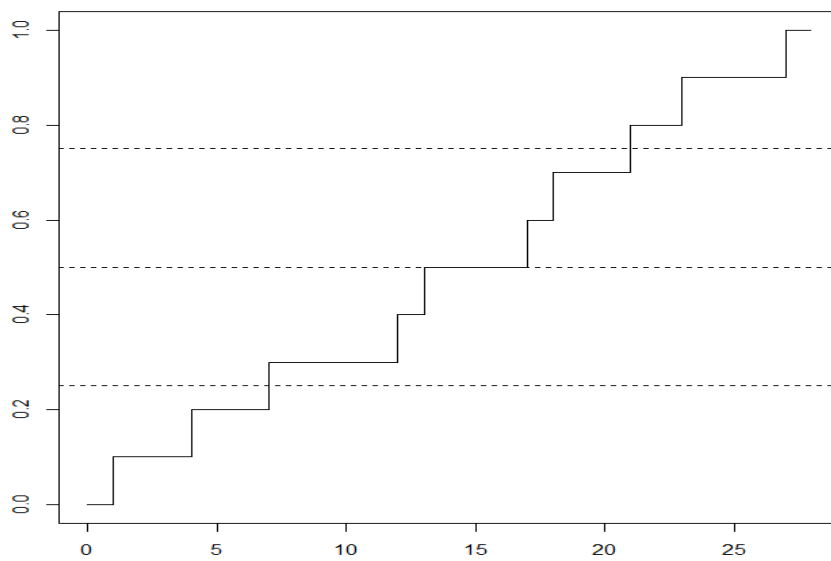
Données triées par ordre croissant

1 4 7 12 13 17 18 21 23 27

Aucune valeur n'est censurée et $k=10$ (toutes les valeurs sont distinctes).

Fonction de répartition

i	$X_{(i)}$	N_i	R_i	$F(X_{(i)})$
0	$-\infty$	0	0	0.0
1	1	1	1	0.1
2	4	2	1	0.2
3	7	3	1	0.3
4	12	4	1	0.4
5	13	5	1	0.5
6	17	6	1	0.6
7	18	7	1	0.7
8	21	8	1	0.8
9	23	9	1	0.9
10	27	10	1	1.0



Percentiles

$$P_{75} = 21$$

$$P_{50} = 15$$

$$P_{25} = 7$$

Par l'ancienne méthode, on retrouve exactement les mêmes résultats.

5 RAPPORT GLOBAL

Dans le rapport global après chaque enquête, on fournit la médiane M, l'écart-type SD et le coefficient de variation CV. On distingue les situations suivantes.

Cas 1: Les quartiles P_{25} , P_{50} et P_{75} existent. On calcule $M = P_{50}$, $SD = 0.74 \times (P_{75} - P_{25})$ et $CV = SD/M$ ($\times 100\%$) comme précédemment.

Cas 2: Le quartile P_{25} n'existe pas mais P_{50} existe. Alors, $M = P_{50}$, $SD = 1.48 \times (P_{75} - P_{50})$ et $CV = SD/M$ ($\times 100\%$) comme précédemment.

Cas 3: Le quartile P_{50} (et donc P_{25}) n'existe pas. Alors, on ne peut pas calculer M et SD. On n'imprime rien sur le rapport (comme c'est aussi déjà le cas lorsque $n < 6$).

6 RAPPORT INDIVIDUEL

Le rapport individuel présente pour chaque test, la valeur du résultat R, la médiane M, l'écart-type SD et la distance (ou Z-score) $D = (R - M)/SD$, non seulement pour l'ensemble des laboratoires mais aussi pour les laboratoires utilisant la même méthode analytique. Un graphique reprenant les limites inférieures (LOF, LIF) et supérieures (UOF, UIF) positionne le résultat R. Pour rappel,

$$LOF = P_{25} - 3H$$

$$LIF = P_{25} - 1.5H$$

$$UIF = P_{75} + 1.5H$$

$$UOF = P_{75} + 3H$$

où $H = P_{75} - P_{25}$.

Il est proposé d'appliquer la présentation suivante:

Cas 1: M et SD n'existent pas. Dans ce cas, on donne le résultat (censuré ou non) mais on ne calcule pas le Z-score (ou distance D) et il n'y a pas de représentation graphique.

Cas 2: M et SD existent et $M - 3 \times SD > 0$. Désignons par R^* la valeur sans le signe "<" du résultat R lorsque celui-ci est censuré. Par exemple, si R vaut <0.5, alors $R^* = 0.5$. Lorsque R n'est pas censuré, $R^* = R$ bien évidemment. On calcule le Z-score (ou distance D), soit

$$D^* = (R^* - M)/SD$$

Si R est censuré, on réaffecte le signe "<" à D^* et on note D la valeur censurée ainsi obtenue ($<D^*$). Sinon, $D = D^*$. On imprime M, SD, R et D (censuré ou non).

Pour la représentation graphique, on procède comme précédemment en localisant la valeur de R sur l'intervalle (LOF – UOF).

Cas 3: M et SD existent mais $M - 3 \times SD \leq 0$. Dans ce cas, si $R^* < M$, on ne calcule pas D! Par contre, si $R^* \geq M$, on procède comme au cas 2.

7 RAPPORT INDIVIDUEL ANNUEL

En fin d'exercice, on calcule les Z-scores et U-scores pour tous les résultats fournis par le laboratoire et on comptabilise la proportion P_z de Z-scores et P_u de U-scores "inacceptables".

Il faut se rappeler qu'un nombre élevé de valeurs censurées dans l'échantillon des résultats fournis par les laboratoires signe la présence d'une concentration faible (voire nulle) dans le spécimen contrôle EEQ. Ceci signifie que les valeurs très basses de Z-scores ou U-scores sont parfaitement acceptables. Il faut donc s'assurer que la concentration du constituant dans le spécimen biologique est pratiquement nulle. Pour ce faire, on calcule la quantité $M - 3 \times SD$. Si $M - 3 \times SD \leq 0$, alors on ne se préoccupe pas des résultats (censurés ou non) inférieurs à M .

En conclusion, pour le rapport annuel, on procède comme suit:

Cas 1: M et SD n'existent pas. On ne calcule ni Z-score, ni U-score. C'est déjà le cas lorsque $n < 6$.

Cas 2: M et SD existent et $M - 3 \times SD > 0$. En posant comme précédemment R^* la valeur censurée R sans le signe "<", on calcule

$$Z^* = \frac{R^* - M}{SD}$$
$$U^* = \frac{R^* - M}{M} (\times 100\%)$$

et on procède comme auparavant. En clair, la valeur R est considérée comme "anormale" si $|Z^*| \geq 3$ ou si $|U^*| \geq d$, où " d " est la limite d'acceptabilité biologique. On imprime "<Z*" (si R est censuré) ou "Z*" (si R est non censuré) sur le rapport et on procède de même avec les U-scores.

Cas 3: M et SD existent mais $M - 3 \times SD \leq 0$. On procède comme au cas 2, sauf si $R^* < M$, auquel cas on ne calcule ni Z-score, ni U-score.

8 COMPARAISON DES DEUX METHODES

Exemple 1 – Digoxine

Considérons les résultats de digoxine (nmol/L) pour l'échantillon A de l'enquête 2 de l'exercice EEQ 2000.

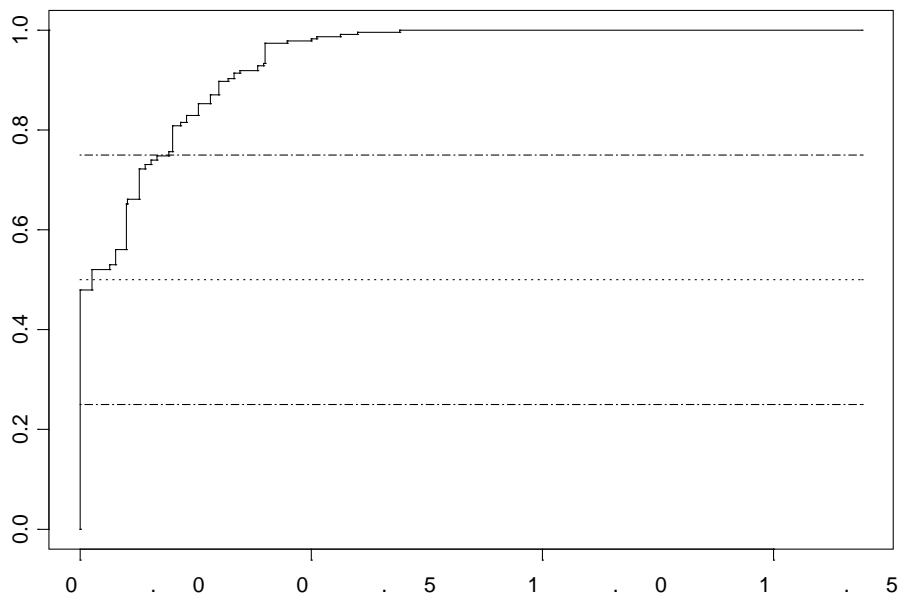
Données triées par ordre croissant

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0.0128 *	0.0256 *	0.0256
0.0256	0.0256	0.0256	0.0512 *	0.064	0.0768
0.0768	0.0768	0.1 *	0.1 *	0.1 *	0.1 *
0.1 *	0.1	0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1	0.1	0.1024
0.128 *	0.128 *	0.128 *	0.128 *	0.128	0.128
0.128	0.128	0.128	0.128	0.128	0.1408
0.1536	0.1664	0.1792 *	0.192 *	0.192 *	0.192 *
0.192 *	0.192	0.2 *	0.2 *	0.2 *	0.2 *
0.2 *	0.2 *	0.2 *	0.2 *	0.2 *	0.2 *
0.2	0.2	0.2	0.2	0.2	0.2
0.2	0.2176 *	0.2176 *	0.2176 *	0.2176 *	0.2176 *
0.2176 *	0.2176 *	0.2176 *	0.2176	0.2304	0.2304
0.2432 *	0.256 *	0.256 *	0.256 *	0.256 *	0.256 *
0.256 *	0.256 *	0.256 *	0.256 *	0.256 *	0.256 *
0.256 *	0.256 *	0.256 *	0.256 *	0.256 *	0.256 *
0.256 *	0.256 *	0.256	0.256	0.256	0.256
0.2816	0.2816	0.2816	0.3 *	0.3 *	0.3 *
0.3 *	0.3 *	0.3 *	0.3 *	0.3 *	0.3 *
0.3 *	0.3 *	0.3 *	0.3 *	0.3 *	0.3
0.3	0.3	0.3	0.3	0.32 *	0.32
0.3328	0.3328	0.3456	0.384 *	0.384 *	0.384 *
0.384 *	0.384 *	0.384 *	0.384 *	0.384 *	0.384 *
0.384 *	0.384 *	0.384 *	0.384 *	0.384 *	0.384 *
0.384 *	0.384 *	0.384 *	0.384 *	0.384 *	0.384 *
0.384 *	0.384	0.384	0.3968	0.4 *	0.4 *
0.4 *	0.4 *	0.4 *	0.4 *	0.4 *	0.4 *
0.4 *	0.4 *	0.4 *	0.4 *	0.4	0.4
0.4	0.4	0.4	0.4	0.4	0.4
0.4	0.448	0.5 *	0.5 *	0.5	0.512 *
0.512 *	0.512 *	0.512 *	0.512 *	0.512 *	0.512
0.5632	0.6 *	0.6 *	0.6 *	0.6 *	0.6 *
0.6	0.64 *	0.64 *	0.64 *	0.6912	

Les données censurées sont suivies d'un astérisque. On constate que sur les 239 valeurs, 120 sont censurées et 119 sont non censurées.

Nouvelle méthode

La nouvelle méthode conduit à la fonction de répartition reprise à la figure ci-dessous.



Le calcul des quartiles fournit les résultats suivants:

$$P_{25} = 0$$

$$P_{50} = 0.0256$$

$$P_{75} = 0.192$$

Dès lors, $SD = 0.142$, et on constate que $M - 3 \times SD = -0.40 < 0$. On peut donc suspecter l'absence de digoxine dans l'échantillon.

Sur la base du Z-score, tout résultat au-delà de $M + 3 \times SD = 0.45$ peut être considéré comme trop élevé. Au vu des données ci-dessus, on est amené à rejeter les 21 résultats les plus élevés de l'échantillon. A titre d'illustration, la valeur censurée < 0.5 doit être rejetée car elle peut valoir 0.49!

Ancienne méthode

L'ancienne méthode, qui n'utilise que les 119 observations non censurées, donne les résultats suivants:

$$P_{25} = 0$$

$$P_{50} = 0.10$$

$$P_{75} = 0.2688$$

Dès lors, $SD = 0.20$ et on constate que $M - 3 \times SD = -0.50$ est également négatif. Toutefois, la limite supérieure $M + 3 \times SD = 0.70$ conduit à l'acceptation de toutes les valeurs de l'échantillon, y compris 0.69!

Rapport individuel

A titre d'exemple, si un laboratoire fournit le résultat <0.5, on a $R^* = 0.5$ et $Z^* = (0.5 - 0.026)/0.142 = 3.34$. Dès lors, le Z-score associé au résultat <0.5 s'écrit <3.34. Il doit être considéré comme "inacceptable" puisque supérieur à 3!

Notons que le U-score n'a pas beaucoup de sens puisque

$$U^* = (0.5 - 0.026)/0.026 \quad \text{soit } 1823\%$$

Exemple 2 – Théophylline

Considérons les résultats de théophylline ($\mu\text{mol/L}$) pour l'échantillon A de l'enquête 2 de l'exercice EEQ 2000.

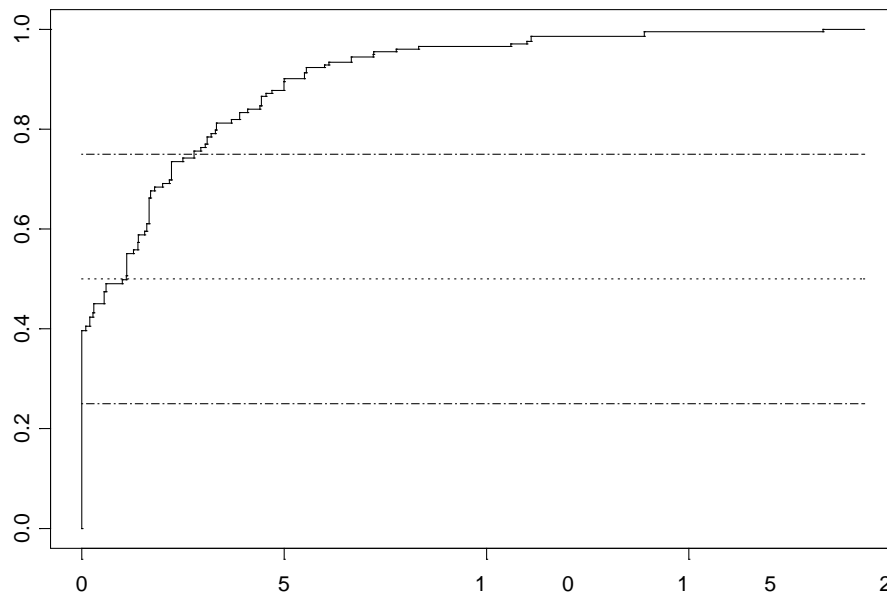
Résultats triés par ordre croissant

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0.1 *	0.1	0.2	0.2	0.2775
0.3	0.3	0.4 *	0.5 *	0.5 *	0.5 *
0.555 *	0.555 *	0.555	0.555	0.555	0.6
0.6	1 *	1	1.1	1.11 *	1.11 *
1.11 *	1.11 *	1.11	1.11	1.11	1.11
1.11	1.11	1.2765	1.3875	1.3875	1.4
1.4	1.554	1.6095	1.6095	1.665 *	1.665
1.665	1.665	1.665	1.665	1.665	1.665
1.7	1.7	1.8	1.998 *	1.998	2.1645
2.22	2.22	2.22	2.22	2.22	2.5
2.775 *	2.775 *	2.775 *	2.775 *	2.775	2.775
2.9415	3.0525	3.1	3.1	3.2	3.3
3.33	3.33	3.7	3.9	3.9	4.1
4.4 *	4.4 *	4.4 *	4.4 *	4.4	4.44 *
4.44 *	4.44 *	4.44 *	4.44 *	4.44 *	4.44 *
4.44 *	4.44 *	4.44	4.44	4.44	4.5 *
4.5 *	4.551 *	4.551 *	4.551 *	4.551 *	4.551 *
4.551 *	4.551 *	4.551 *	4.551	4.7	4.995
4.995	4.995	5	5.5 *	5.5 *	5.5
5.5	5.55 *	5.55 *	5.55 *	5.55 *	5.55 *
5.55 *	5.55 *	5.55 *	5.55 *	5.55 *	5.55 *
5.55 *	5.55 *	5.55 *	5.55	5.55	5.6 *
5.6 *	6 *	6	6.105	6.66	6.66
7.2	7.215	7.77	8.325	8.9 *	10 *
10.6	11	11.1 *	11.1 *	11.1 *	11.1
11.1	13.8 *	13.875 *	13.875 *	13.875 *	13.875 *
13.875 *	13.875 *	13.875 *	13.875 *	13.875 *	13.875 *
13.875 *	13.9 *	13.9 *	13.9 *	13.9 *	13.9 *
13.9 *	13.9 *	13.9 *	13.9 *	13.9 *	13.9 *
13.9	13.9	14 *	14.208 *	15 *	16.65 *
18.315					

Les données censurées sont marquées d'un astérisque. On constate qu'il y a 92 valeurs censurées et 131 valeurs non censurées.

Nouvelle méthode

La fonction de répartition est reprise à la figure ci-dessous.



Les quartiles calculés ont pour valeur:

$$P_{25} = 0$$

$$P_{50} = 1.10$$

$$P_{75} = 2.775$$

Dès lors, $SD = 2.05$ et $M - 3 \times SD = -5.06!$ On peut considérer que l'échantillon ne contient pas de théophylline.

Puisque $M + 3 \times SD = 7.26$, on constate que 41 (18%) valeurs de l'échantillon sont trop élevées et dès lors en-dehors des normes.

Ancienne méthode

L'ancienne méthode conduit aux résultats suivants:

$$P_{25} = 0$$

$$P_{50} = 1.40$$

$$P_{75} = 3.52$$

Dès lors, $SD = 2.60$ et $M - 3 \times SD = -6.41 < 0$. Enfin, la limite supérieure $M + 3 \times SD = 9.20$ conduit à rejeter 38 valeurs au lieu de 41. En effet, les valeurs 7.77, 8.33 et <8.9 sont à présent acceptées!

9 PROBLEME DES PETITS GROUPES

Les analyses statistiques effectuées sur des petits groupes de données EEQ (par exemple pour une méthode analytique donnée) montrent qu'il faut être prudent dans l'interprétation des résultats.

Il peut arriver que toutes les données soient censurées. Par exemple, Digoxine (Echantillon A, EEQ 2000/2) pour la méthode "*Non isotopic – DPC – immulite*" ($n=6$):

<0.1 <0.5 <0.32 <0.60 <0.64 <0.64

Dans ce cas, on ne peut évidemment rien faire!

Il peut arriver qu'une majorité de données soient nulles. Par exemple, Digoxine (Echantillon A, EEQ 2000/2) pour la méthode "*Non-isotopic – Dade Behring – Emit*" ($n=16$).

0	0	0	0	0	0	0	0
0	<0.0256	0.1	0.128	<0.1792	<0.192	<0.3	<0.6

Dans ce cas $P_{25} = P_{50} = P_{75} = 0$ et donc $M = 0$ et $SD = 0$. On ne peut rien faire!

10 RECOMMANDATIONS FINALES

Pour un spécimen contrôle EEQ dont la concentration d'un constituant est très faible, voire nulle, il est recommandé de

- ne pas procéder à une interprétation des résultats par méthode analytique car elle est peu fiable
- de réaliser une interprétation globale, en regroupant l'ensemble des laboratoires
- d'utiliser la méthode des Z-scores uniquement pour les valeurs élevées
- d'éviter la méthode des U-scores qui n'a plus guère de sens dans la mesure où le dénominateur dans la formule est proche de 0.

Enfin, il est recommandé d'utiliser la nouvelle méthode qui tient compte de l'ensemble des résultats, censurés ou non, et qui est en général plus rigoureuse que l'ancienne comme le montre la plupart des exemples.