Use for training only.

Annex C
(informative)

Robust analysis

C.1 General

Interlaboratory comparisons present unique challenges for data analysis. While most interlaboratory comparisons provide unimodal and approximately symmetric data, most proficiency testing data sets include a proportion of results that are unexpectedly distant from the majority. These can arise for a variety of reasons; for example, from less experienced participants, from less precise, or perhaps new, measurement methods, or from participants who did not understand the instructions or who processed the proficiency test items incorrectly. Such outlying results can be highly variable and make conventional statistical techniques, including the mean and standard deviation, unreliable.

It is recommended (see 6.5.1) that proficiency testing providers use statistical techniques that are robust to outliers. Many such techniques have been proposed in the statistical literature, and many of those have been used successfully for proficiency testing. Most robust techniques additionally confer resistance to asymmetric outlier distributions.

This Annex describes several techniques that have been applied in proficiency testing and have different capabilities regarding robustness to contaminated populations (for example efficiency and breakdown point), and differing simplicity of application. They are presented here in order of simplicity (simplest first, most complex last), which is approximately inversely related to efficiency because the more complex estimators tend to have been developed in order to improve efficiency.

NOTE 1 Annex D provides further information on efficiency, breakdown point and sensitivity to minor modes - three important indicators of the performance of various robust estimators.

NOTE 2 Robustness is a property of the estimation algorithm, not of the estimates it produces, so it is not strictly correct to call the averages and standard deviations calculated by such an algorithm "robust". However, to avoid the use of excessively cumbersome terminology, the terms "robust average" and "robust standard deviation" are understood in this document to mean estimates of the population mean or of the population standard deviation calculated using a robust algorithm.

C.2 Simple outlier-resistant estimators for the population mean and standard deviation

C.2.1 The median

The median is a simple and highly outlier-resistant estimator of the population mean for symmetric distributions. To determine the median, denoted $med(x)$:

a) Denote the p items of data, sorted into increasing order, by

$x_{\{1\}}, x_{\{2\}}, ..., x_{\{p\}}$

b) calculate

© ISO 2022 – All rights reserved

Use for training only.

$$med(x) = \begin{cases} x_{\{(p+1)/2\}} & p \quad \text{odd} \\ \dfrac{\left[x_{\{p/2\}} + x_{\{1+p/2\}}\right]}{2} & p \quad \text{even} \end{cases} \tag{C.1}$$

C.2.2 Scaled median absolute deviation MAD_e

The scaled median absolute deviation $MAD_e(x)$ provides an estimate of the population standard deviation for normally distributed data and is highly resistant to outliers. To calculate $MAD_e(x)$:

a) Calculate the absolute differences d_i (for i = 1 to p) from

$$d_i = \left|x_i - med(x)\right| \tag{C.2}$$

b) Calculate $MAD_e(x)$ from

$$MAD_e(x) = 1{,}483\ med(d) \tag{C.3}$$

If 50 % or more of the participant results are the same, then $MAD_e(x)$ will be zero, and it may be necessary to use the nIQR in C.2.3, an arithmetic standard deviation (after outlier removal), or the procedure described in C.5.2.

C.2.3 Normalized interquartile range nIQR

A robust estimator of the standard deviation similar to $MAD_e(x)$ and slightly simpler to obtain has proved to be useful in many proficiency testing schemes, and can be obtained from the difference between the 75th percentile (or 3rd quartile) and 25th percentile (or 1st quartile) of the participant results. This statistic is commonly called the 'normalized InterQuartile Range' (or nIQR), and it is calculated as in Formula (C.4):

$$\mathrm{nIQR}(x) = 0{,}7413(Q_3(x) - Q_1(x)) \tag{C.4}$$

where

$Q_1(x)$ denotes the 25th percentile of x_i (i = 1,2,...,p)

$Q_3(x)$ denotes the 75th percentile of x_i (i = 1,2,...,p)

If the 75th and 25th percentiles are the same, the nIQR will be zero (as will $MAD_e(x)$) and an alternative procedure such as an arithmetic standard deviation (after outlier removal) or the procedure at C.5.2 should be used to calculate the robust standard deviation.

NOTE 1 The nIQR only requires sorting the data once compared to MAD_e but has breakdown point of 25 % (see Annex D), while MAD_e has breakdown point of 50 %. MAD_e can therefore tolerate an appreciably higher proportion of outliers than nIQR.

NOTE 2 Both nIQR and the MAD_e estimators show appreciable negative bias at $p<30$ which can adversely affect scores if these estimates are used in scoring participant results.

NOTE 3 Different statistical packages can use different algorithms for calculating quartiles, and can therefore produce slightly different nIQR.

NOTE 4 An example using simple robust estimators is included in E.3.

© ISO 2022 – All rights reserved

Use for training only.

C.3 Robust analysis: Algorithm A

C.3.1 Algorithm A with iterated scale

This algorithm yields robust estimates of the mean and standard deviation of the data to which it is applied.

Denote the p items of data, sorted into increasing order, by:

$$x_{\{1\}}, x_{\{2\}}, \ldots, x_{\{p\}}$$

Denote the robust average and robust standard deviation of these data by x^* and s^*.

Calculate initial values for x^* and s^* as:

$$x^* = \text{median of } x_i \ (i = 1, 2, \ldots, p) \tag{C.5}$$

$$s^* = 1{,}483 \text{ median of } \left| x_i - x^* \right| \text{ with } (i = 1, 2, \ldots, p) \tag{C.6}$$

NOTE 1 Algorithms A and S given in this annex are reproduced from ISO 5725-5, with a slight addition to Algorithm A to specify a stopping criterion: no change in the 3rd significant figures of the robust mean and standard deviation.

NOTE 2 In some cases more than half of the results x_i will be identical (for example thread count in fabric, or electrolytes in serum). In these cases the initial value of s^* will be zero and the robust procedure will not perform correctly. In the case that the initial $s^* = 0$, it is acceptable to substitute the sample standard deviation, after checking for any gross outliers that could make the sample standard deviation unreasonably large. This substitution is made only for the initial s^*, and after that the iterative algorithm can proceed as described.

Update the values of x^* and s^* as follows. Calculate:

$$\delta = 1{,}5 s^* \tag{C.7}$$

For each x_i ($i = 1, 2, \ldots, p$), calculate:

$$x_i^* = \begin{cases} x^* - \delta & \text{when} \quad x_i < x^* - \delta \\ x^* + \delta & \text{when} \quad x_i > x^* + \delta \\ x_i & \text{otherwise} \end{cases} \tag{C.8}$$

Calculate the new values of x^* and s^* from:

$$x^* = \sum_{i=1}^{p} x_i^* / p \tag{C.9}$$

$$s^* = 1{,}134 \sqrt{\sum_{i=1}^{p} \left(x_i^* - x^* \right)^2 / (p - 1)} \tag{C.10}$$

where the summation is over i.

The robust estimates x^* and s^* may be derived by an iterative calculation, i.e. by updating the values of x^* and s^* several times using Formulae C.7 to C.10, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figures of the robust mean and robust standard deviation (x^* and s^*). Alternative convergence criteria can be determined according to the design and reporting requirements for proficiency test results.

NOTE 3 Examples of use of Algorithm A with iterated scale are provided in E.1 and E.3.

 © ISO 2022 – All rights reserved

Use for training only.

C.3.2 Variants of Algorithm A

Algorithm A with iterated scale in C.3.1 has modest breakdown (approximately 25 % for large data sets[25]) and the starting point for s^* suggested in C.3.1 for data sets where $\mathrm{MAD_e}(x)$ is zero can seriously degrade outlier resistance when there are severe outliers in the data set. The following variations should be considered where the proportion of outliers is expected to be over 20 % in any data set or where the initial value for s^* is adversely affected by extreme outliers:

a) Replace $\mathrm{MAD_e}$ with $med\left(\left|x_i - \bar{x}\right|\right)$ when $\mathrm{MAD_e}$=0, or use an alternative estimator such as that described in C.5.1 or the arithmetic standard deviation (after outlier removal).

b) Where the robust standard deviation is not used in scoring, use $\mathrm{MAD_e}$ (amended as i) above) and do not update s^* during iteration. Where the robust standard deviation is used in scoring, replace s^* with the Q estimator described in C.5 and do not update s^* during iteration.

NOTE Variant b) improves the breakdown point of Algorithm A to 50 %[25], allowing the algorithm to cope with a higher proportion of outliers.

C.4 Robust analysis: Algorithm S

This algorithm is applied to standard deviations (or ranges), which are calculated when participants submit m replicate results for a measurand in a proficiency test item, or in a study with m identical proficiency test items. It yields a robust pooled value of the standard deviations or ranges to which it is applied.

Denote the p standard deviations or ranges, sorted into increasing order, by:

$$w_{\{1\}}, w_{\{2\}}, ..., w_{\{p\}}$$

Denote the robust pooled value by w^*, and the degrees of freedom associated with each w_i by ν. (When w_i is a range, $\nu = 1$. When w_i is the standard deviation of m test results, $\nu = m - 1$.) Obtain the values of ξ and η required by the algorithm from Table C.1.

Calculate an initial value for w^* as:

$$w^* = \text{median of } w_i \ (i = 1, 2, ..., p) \tag{C.11}$$

NOTE If more than half of the w_i are zero then the initial w^* will be zero and the robust procedure will not perform correctly. When the initial w^* is zero, substitute the arithmetic pooled average standard deviation (or average range) after eliminating any extreme outliers that can influence the average. This substitution is only for the initial w^*, after which the procedure is continued as described.

© ISO 2022 – All rights reserved

Use for training only.

Update the value of w^* as follows. Calculate:

$$\psi = \eta \times w^* \tag{C.12}$$

For each w_i (i = 1, 2, ..., p), calculate:

$$w_i^* = \begin{cases} \psi & \text{if } w_i > \psi \\ w_i & \text{otherwise} \end{cases} \tag{C.13}$$

Calculate the new value of w^* from:

$$w^* = \xi \sqrt{\sum_{i=1}^{p} \left(w_i^*\right)^2 / p} \tag{C.14}$$

The robust estimate w^* is calculated by an iterative calculation by updating the value of w^* several times, until the process converges. Convergence may be assumed when there is no change from one iteration to the next in the third significant figure of the robust estimate.

NOTE Algorithm S provides an estimate of the population standard deviation when supplied with standard deviations from a single normal distribution, and hence provides an estimate of the repeatability standard deviation when the assumptions of ISO 5725-2 apply.

Table C.1 — Factors required for robust analysis: Algorithm S

Degrees of freedom ν	Limit factor η	Adjustment factor ξ
1	1,645	1,097
2	1,517	1,054
3	1,444	1,039
4	1,395	1,032
5	1,359	1,027
6	1,332	1,024
7	1,310	1,021
8	1,292	1,019
9	1,277	1,018
10	1,264	1,017
NOTE The values of ξ and η are derived in ISO 5725-5:1998, Annex B.		

C.5 Computationally intensive robust estimators: *Q* method and Hampel estimator

C.5.1 Rationale for computationally intensive estimators

The robust estimators of the population mean and standard deviation described in C.2 and C.3 are useful when computational resources are limited, or when it is necessary to provide concise explanations of the statistical procedures. These procedures have proven to be useful in a wide variety of situations, including for proficiency testing schemes in new areas of testing or calibration and in economies where proficiency testing has not previously been available. However, these techniques can become unreliable when more than 20 % of results are outliers, or where there are bimodal (or multimodal) distributions, and some may become unacceptably variable for smaller numbers of participants. Further, none can handle replicated data from participants. ISO/IEC 17043 requires that these situations will be anticipated by design or will be detected by competent review prior to performance evaluation, but there are occasions when this may not be possible.

 © ISO 2022 – All rights reserved

Use for training only.

In addition, some of the robust techniques described in C.2 and C.3 are lacking in terms of statistical efficiency - if the number of participants is less than 50, and the robust mean and/or standard deviation are used for scoring there is a considerable risk for misclassifying participants due to the use of ineffective statistical methods.

Robust techniques that combine good efficiency (that is, comparatively low variability) with tolerance for a high proportion of outliers tend to be more complex and require more computational resources, but the techniques are referenced in available literature and International Standards. Some of these additionally provide useful performance gains when the underlying distribution of data is skewed or when some results are quoted as below a detection or reporting limit.

The following paragraphs describe some high-efficiency, high-breakdown methods for estimating standard deviation and location (mean) that are useful for data with larger proportions of outliers and that show lower variability than simpler estimators. One of the estimators described can also be used to estimate a reproducibility standard deviation when participants report multiple observations.

C.5.2 Determination of a robust standard deviation using Q and Q_n methods

C.5.2.1 Q_n [34] is a high-breakdown, high-efficiency estimator of the population standard deviation which is unbiased for normally distributed data (that is, under the assumption that there are no outliers). Q_n uses a single reported result (including a mean or median of replicates) for each participant. The calculation relies on the use of pairwise differences within the data set and therefore it is not dependent on an estimate of the mean or median of the data. The implementation described here includes corrections to ensure that the estimate is unbiased for all practical data set sizes.

To calculate Q_n for a data set $(x_1, x_2, \ldots x_p)$ with p reported results:

a) Calculate the $p(p\text{-}1)/2$ absolute differences

$$d_{ij} = \left|x_i - x_j\right| \text{ for } i = 1,2 \ldots p-1 \text{ and } j = i+1, i+2 \ldots p \qquad \text{(C.15)}$$

b) Denote the ordered differences d_{ij} by

$$d_{\{1\}}, d_{\{2\}} \ldots d_{\{p(p\text{-}1)/2\}} \qquad \text{(C.16)}$$

c) Calculate

$$k = \frac{h(h-1)}{2} \qquad \text{(C.17)}$$

that is, k is the number of distinct pairs chosen from h objects, where:

$$h = \begin{cases} p/2 & p \quad \text{even} \\ (p-1)/2 & p \quad \text{odd} \end{cases} \qquad \text{(C.18)}$$

d) Calculate Q_n as

$$Q_\text{n} = 2{,}2219 d_{\{k\}} b_p \qquad \text{(C.19)}$$

where b_p is selected from Table C.2 for a particular number p of data points or, for $p > 12$, is calculated from

$$b_p = \frac{1}{r_p + 1} \qquad \text{(C.20)}$$

where

© ISO 2022 – All rights reserved

Use for training only.

$$r_p = \begin{cases} \frac{1}{p}\left[1,6019 + \frac{1}{p}\left(-2,128 - \frac{5,172}{p}\right)\right] & p \quad \text{odd} \\ \frac{1}{p}\left[3,6756 + \frac{1}{p}\left(1,965 + \frac{1}{p}\left(6,987 - \frac{77}{p}\right)\right)\right] & p \quad \text{even} \end{cases} \tag{C.21}$$

NOTE 1 The factor of 2,2219 is a correction factor to give an unbiased estimate of standard deviation for large p. The correction factors b_p for small p are in Table C.2 and the calculation for r_p for $p > 12$ are as provided in Reference [41] from extensive simulation and subsequent regression analysis.

NOTE 2 The simple algorithm described above requires considerable computing resources for larger data sets, for example $p > 1000$. A fast and memory-efficient implementation capable of handling much larger data sets has been published with full computer code[42] for use with larger data sets; Reference [42] cited acceptable performance for p over 8 000 at the time of publication.

Table C.2 — Correction factor b_p for $2 \le p \le 12$

p	2	3	4	5	6	7	8	9	10	11	12
b_p	0,3994	0,9937	0,5132	0,8440	0,6122	0,8588	0,6699	0,8734	0,7201	0,8891	0,7574

C.5.2.2 The Q method produces a high-breakdown, high-efficiency estimate of the standard deviation of proficiency testing results reported by different laboratories. The Q method is not only robust against outlying results, but also against a situation where many test results are equal, e.g. due to quantitative data on a discontinuous scale or due to rounding distortions. In such a situation other Q-like methods can fail because many pairwise differences are zero.

The Q method can be used for proficiency testing both with single results per participant (including a mean or median of replicates) and for replicates. The direct use of replicates in the calculation improves the efficiency of the method.

The calculation relies on the use of pairwise differences within the data set and is therefore not dependent on an estimate of the mean or median of the data. The method is known as Q/Hampel when it is used together with the finite step algorithm for the Hampel estimator described in C.5.3.3.

Denote the reported measurement results, grouped by laboratory, by:

$$\underbrace{y_{11},\ldots,y_{1n_1}}_{Lab\ 1},\underbrace{y_{21},\ldots,y_{2n_2}}_{Lab\ 2},\ldots,\underbrace{y_{p1},\ldots,y_{pn_p}}_{Lab\ p} \tag{C.22}$$

Calculate the cumulative distribution function of all absolute between-laboratory differences

$$H_1(x) = \frac{2}{p(p-1)} \sum_{1 \le i < j \le p} \frac{1}{n_i n_j} \sum_{k=1}^{n_i} \sum_{m=1}^{n_j} \mathbf{I}\left\{\left|y_{ik} - y_{jm}\right| \le x\right\} \tag{C.23}$$

where $\mathbf{I}\left\{\left|y_{ik} - y_{jm}\right| \le x\right\} = \begin{cases} 1 & \text{if } \left|y_{ik} - y_{jm}\right| \le x \\ 0 & \text{otherwise} \end{cases}$ denotes the indicator function.

© ISO 2022 – All rights reserved

Use for training only.

Denote the discontinuity points of $H_1(x)$ by:

$x_1, ..., x_r$, where $x_1 < x_2 < ... < x_r$

Calculate for all positive discontinuity points $x_1, ..., x_r$:

$$G_1\left(x_i\right) = \begin{cases} 0,5 \cdot \left(H_1\left(x_i\right) + H_1\left(x_{i-1}\right)\right) & \text{if} \quad i \geq 2 \\ 0,5 \cdot H_1\left(x_1\right) & \text{if} \quad i = 1; x_1 > 0 \end{cases} \tag{C.24}$$

and let

$G_1(0)=0$

Calculate the function $G_1(x)$ for all x out of the interval $[0, x_r]$ by linear interpolation between discontinuity points $0 \le x_1 < x_2 < ... < x_r$.

Calculate the robust standard deviation s^* of test results of different laboratories

$$s^* = \frac{G_1^{-1}\left(0,25 + 0,75 \cdot H_1\left(0\right)\right)}{\sqrt{2}\Phi^{-1}\left(0,625 + 0,375 \cdot H_1\left(0\right)\right)} \tag{C.25}$$

where $H_1(0)$ is calculated as in Formula (C.23) and is equal to zero unless there are exact ties in the data set, and where $\Phi^{-1}(q)$ is the q^{th} quantile of the standard normal distribution.

NOTE 1 This algorithm does not depend on a mean value; it can be used together with either a value from combined participant results or a specified reference value.

NOTE 2 Other variants of the Q method provide robust estimates of both repeatability and reproducibility standard deviation[25,34].

NOTE 3 The theoretical basis for the Q method, including asymptotic performance and finite sample breakdown, are described in References [26] and [34].

NOTE 4 If the underlying data of the participants represent single measurement results obtained with one specific measurement method, the robust standard deviation is an estimate of the reproducibility standard deviation as in Formula (C.21).

NOTE 5 The reproducibility standard deviation is not necessarily the most appropriate standard deviation for use in proficiency testing because it is usually an estimate of the dispersion of single results and not an estimate of the dispersion of means or medians of replicated results from each participant. However the dispersion of means or medians of replicated results is only slightly below the dispersion of single results of different laboratories, if the ratio of reproducibility standard deviation divided by the repeatability standard deviation is greater than 2. If this ratio is below 2, for scoring in proficiency testing it can be useful to replace the reproducibility standard deviation s_R by the corrected value $\sqrt{s_R^2 - \frac{m-1}{m} s_r^2}$, where m denotes number of replicates and s_r^2 the repeatability variance as calculated in Reference [35], or to use not the replicates but the mean of replicates per participant for the Q method.

NOTE 6 Note 5 applies only if the scoring is conducted on the basis of means or medians of replicated results. If the replicates are blind replicate proficiency test items, scores are assumed to be given for each replicate. In this case the reproducibility standard deviation is the most appropriate standard deviation.

NOTE 7 An example to which the Q method has been applied is shown in E.3.

C.5.3 Determination of a robust mean using the Hampel estimator

C.5.3.1 The Hampel estimate is a highly robust and efficient estimate of the overall mean of results reported by different laboratories. As there is no explicit formula for obtaining the Hampel estimate, in this paragraph two algorithms are provided. The first one can be more easily implemented but may lead

© ISO 2022 – All rights reserved

Use for training only.

to deviating results in different implementations. The second one provides unique results depending only on the underlying standard deviation.

C.5.3.2 The following calculation provides an iterative reweighting scheme for obtaining the Hampel estimate of location.

a) Denote the data as $x_1, x_2 \ldots x_p$

b) Set x^* to $med(x)$ (see C.2.1)

c) Set s^* to a suitable robust estimate of standard deviation, for example $\mathrm{MAD_e}$, Q_n or s^* from the Q method.

d) For each data point x_i, calculate q_i from Formula (C.26):

$$q_i = \left| \frac{x_i - x^*}{s^*} \right| \tag{C.26}$$

e) Calculate weights w_i from Formula (C.27)

$$w_i = \begin{cases} 0 & |q| > 4{,}5 \\ (4{,}5 - q)/q & 3 < |q| \le 4{,}5 \\ 1{,}5/q & 1{,}5 < |q| \le 3{,}0 \\ 1 & |q| \le 1{,}5 \end{cases} \tag{C.27}$$

f) Recalculate x^* from Formula (C.28)

$$x^* = \frac{\sum_{i=1}^{p} w_i x_i}{\sum_{i=1}^{p} w_i} \tag{C.28}$$

g) Repeat steps d) to f) until x^* converges. Convergence may be assumed when the change in x^* from one iteration to the next is less than $0{,}01\, s^*/\sqrt{p}$, corresponding to approximately 1 % of the standard error in x^*. Other, more precise, convergence criteria may be used.

This implementation of the Hampel estimator is not guaranteed to have a unique solution or to result in the best solution because a poor choice of initial location x^* and/or s^* may exclude important parts of the data set. The proficiency testing provider should accordingly implement measures to check for the possibility of a poor solution or provide unambiguous rules for choice of location. The most common rule is to choose the solution nearest the median. Reviewing the results to ensure that no large proportion of the data set is outside the range $|q|$>4.5 can also assist in confirming a viable solution.

NOTE 1 This implementation of Hampel's estimator has approximately 96 % efficiency for normally distributed data.

NOTE 2 An example using this implementation is given in E.3

NOTE 3 Hampel's estimator can be tuned for greater efficiency or greater resistance to outliers by changing the weight function. The general form of the weighting function is

© ISO 2022 – All rights reserved

Use for training only.

$$w_i = \begin{cases} 0 & |q| > c \\ a(c-q)/[q(c-b)] & b < |q| \le c \\ a/q & a < |q| \le b \\ 1 & |q| \le a \end{cases}$$

where a, b and c are tuning parameters. For the implementation here, a = 1,5, b = 3,0 and c = 4,5. Greater efficiency is obtained by increasing the range; improved resistance to outliers or minor modes is obtained by reducing the range.

C.5.3.3 The following finite step algorithm yields the Hampel estimate of location without iterative reweighting[25].

Calculate the arithmetic means for each laboratory, now labelled $y_1, y_2, ... y_p$.

Calculate the robust mean, x^*, by solving Formula (C.29):

$$\sum_{i=1}^{p} \Psi\left(\frac{y_i - x^*}{s^*}\right) = 0 \quad \text{(C.29)}$$

where

$$\Psi(q) = \begin{cases} 0 & q \le -4,5 \\ -4,5-q & -4,5 < q \le -3 \\ -1,5 & -3 < q \le -1,5 \\ q & -1,5 < q \le 1,5 \\ 1,5 & 1,5 < q \le 3 \\ 4,5-q & 3 < q \le 4,5 \\ 0 & q > 4,5 \end{cases} \quad \text{(C.30)}$$

and s^* is the robust standard deviation according to the Q method.

The exact solution may be obtained in a finite number of steps, which means not iteratively, using the property that ψ in the argument of x^* is partially linear, bearing in mind that the interpolation nodes on the left side of Formula (C.29) (interpreted here as a function of x^*) are as follows:

Calculate all interpolation nodes

— for the first value y_1:

$$d_1 = y_1 - 4,5s^*,\ d_2 = y_1 - 3s^*,\ d_3 = y_1 - 1,5s^*,\ d_4 = y_1 + 1,5s^*,$$
$$d_5 = y_1 + 3s^*,\ d_6 = y_1 + 4,5s^*$$

— for the second value y_2:

$$d_7 = y_2 - 4,5s^*,\ d_8 = y_2 - 3s^*,\ d_9 = y_2 - 1,5s^*,\ d_{10} = y_2 + 1,5s^*,$$
$$d_{11} = y_2 + 3s^*,\ d_{12} = y_2 + 4,5s^*$$

— and so on for all values $y_3, ..., y_p$.

Sort these data $d_1, d_2, d_3, \ldots, d_{6 \cdot p}$ in ascending order,. $d_{\{1\}}, d_{\{2\}}, d_{\{3\}}, \ldots, d_{\{6 \cdot p\}}$

© ISO 2022 – All rights reserved

Use for training only.

Then calculate for each $m = 1,\ldots,(6 \cdot p - 1)$

$$p_m = \sum_{i=1}^{p} \Psi\left(\frac{y_i - d_{\{m\}}}{s^*}\right)$$

and check whether

a) $p_m = 0$. If so, $d_{\{m\}}$ is a solution of Formula (C.29).

b) $p_{m+1} = 0$. If so, $d_{\{m+1\}}$ is a solution of Formula (C.29).

c) $p_m \cdot p_{m+1} < 0$. If so, $x_m = d_{\{m\}} - \dfrac{p_m}{\dfrac{p_{m+1} - p_m}{d_{(m+1)} - d_{(m)}}}$ is a solution of Formula (C.29).

Let S denote the set of all of these solutions of Formula (C.29).

The solution $x^* \in S$ nearest the median is used as location parameter x^*, i.e.

$$\left|x^* - med\left(y_1, y_2, \ldots, y_p\right)\right| = \min\left\{\left|x - med\left(y_1, y_2, \ldots, y_p\right)\right|; x \in S\right\}$$

Several solutions may exist. If there are two solutions nearest the median, or if there is no solution at all, the median itself is used as location parameter x^*.

NOTE 1 This implementation of Hampel's estimator has approximately 96 % efficiency for normally distributed data.

NOTE 2 If this estimation method is used, laboratory results differing from the mean by more than 4,5 times the reproducibility standard deviation no longer have any effect on the calculation result, i.e. they are treated as outliers.

C.5.4 The *Q*/Hampel method

The method known as *Q*/Hampel uses the *Q* method described in C.5.3.2 for the calculation of the robust standard deviation s^* together with the finite step algorithm for the Hampel estimator described in C.5.3.3 for the calculation of the location parameter x^*.

When participants report multiple observations, the *Q* method described in C.5.3.2 is used for the calculation of the robust reproducibility standard deviation s_R. For the calculation of the robust repeatability standard deviation s_r a second algorithm using the pairwise differences within the laboratories is applied.

NOTE A web application for the *Q*/Hampel method is available[37].

C.6 Other robust techniques

The methods described in this Annex do not constitute a comprehensive collection of valid approaches, and none is guaranteed to be optimal for all situations. Other robust estimators may be used at the discretion of the proficiency testing provider, subject to demonstration, by reference to known efficiency, breakdown point and any other appropriate properties, that they fulfil the particular requirements of the proficiency testing scheme.

© ISO 2022 – All rights reserved