INTERNATIONAL STANDARD

ISO 4259-1

First edition
2017-11

Petroleum and related products — Precision of measurement methods and results —

Part 1: Determination of precision data in relation to methods of test

Produits pétroliers — Fidélité des méthodes de mesure et des résultats —

Partie 1: Détermination des valeurs de fidélité relatives aux méthodes d'essai

Reference number
ISO 4259-1:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 28, *Petroleum and related products, fuels and lubricants from natural or synthetic sources.*

This first edition of ISO 4259-1, together with ISO 4259-2, cancels and replaces ISO 4259, which has been technically revised.

A list of all parts in the ISO 4259 series can be found on the ISO website.

Introduction

For purposes of quality control and to check compliance with specifications, the properties of commercial petroleum products are assessed by standard laboratory test methods. Two or more measurements of the same property of a specific sample by a specific test method, or, by different test methods that purport to measure the same property, will not usually give exactly the same result. It is, therefore, necessary to take proper account of this fact, by arriving at statistically based estimates of the precision for a method, i.e. an objective measure of the degree of agreement expected between two or more results obtained in specified circumstances.

This document makes reference to ISO 3534-2[1], which gives a different definition of true value (see 3.23). This document also refers to ISO 5725-2. The latter is required in particular and unusual circumstances (see 5.3.1) for the purpose of estimating precision.

The two parts of ISO 4259 encompass both the derivation of precision estimates and the application of precision data. They combine the information in ASTM D6300[2] regarding the determination of the precision estimates and the information in ASTM D3244[3] for the utilization of test data.

A glossary of the variables used in this document and ISO 4259-2 is included as Annex I in this document.

Petroleum and related products — Precision of measurement methods and results —

Part 1: Determination of precision data in relation to methods of test

1 Scope

This document specifies the methodology for the design of an Interlaboratory Study (ILS) and calculation of precision estimates of a test method specified by the study. In particular, it defines the relevant statistical terms (Clause 3), the procedures to be adopted in the planning of ILS to determine the precision of a test method (Clause 4), and the method of calculating the precision from the results of such a study (Clauses 5 and 6).

The procedures in this document have been designed specifically for petroleum and petroleum related products, which are normally considered as homogeneous. However, the procedures described in this document can also be applied to other types of homogeneous products. Careful investigations are necessary before applying this document to products for which the assumption of homogeneity can be questioned.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 5725-2, *Accuracy (trueness and precision) of measurement methods and results — Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
analysis of variance
ANOVA
technique that enables the total variance of a method to be broken down into its component factors

3.2
accepted reference value
ARV
agreed-upon reference value for a specific property of a material determined using an accepted reference method and protocol, e.g. derived from an ILS

3.3
between laboratory variance
component of the total variance attributable to the difference between the means of different laboratories

Note 1 to entry: When results obtained by more than one laboratory are compared, the scatter is usually wider than when the same number of tests is carried out by a single laboratory, and there is some variation between means obtained by different laboratories. These give rise to the between laboratory variance which is that component of the overall variance due to the difference in the means obtained by different laboratories.

Note 2 to entry: There is a corresponding definition for between operator variance.

Note 3 to entry: The term "between laboratory" is often shortened to "laboratory" when used to qualify representative parameters of the dispersion of the population of results, for example as "laboratory variance".

3.4
bias
<of a test method> difference between the population mean of test results from a very large number of different laboratories for the property of a material obtained using a specific test method versus the accepted reference value for the property where this is available

Note 1 to entry: See Note 1 to entry in 3.13 for an interpretation of "population mean of test results".

3.5
blind coding
assignment of a different number to each sample so that no other identification or information on any sample is given to the operator

3.6
check sample
sample taken at the place where a product is exchanged, i.e. where the responsibility for the product quality passes from the supplier to the recipient

3.7
degrees of freedom
divisor used in the calculation of variance

Note 1 to entry: The definition applies strictly only in the simplest cases. Definitions for more complex cases are beyond the scope of this document.

3.8
determination
process of carrying out the series of operations specified in a test method, whereby a single value is obtained

3.9
interlaboratory study
ILS
study specifically designed to estimate the repeatability and reproducibility of a standard test method achieved at a fixed point in time by multiple laboratories through the statistical analysis of their test results obtained on aliquots prepared from multiple materials

3.10
known value
quantitative value for a property that can be theoretically derived or calculated by the preparation of the sample

Note 1 to entry: The known value does not always exist, for example for empirical tests such as flash point.

3.11
mean
sum of a set of results divided by the number of results

3.12
mean square
sum of squares divided by the degrees of freedom

3.13
normal distribution
probability distribution of a continuous random variable, x, such that, if x is any real number, the probability density is as shown in Formula (1):

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],-\infty<x<\infty \quad (1)$$

Note 1 to entry: In the context of modelling a distribution of test results, μ is the population mean, or true value (see 3.23) of the property as determined by a specific test method; σ is the standard deviation of the normal distribution used to describe the distribution of an infinite number of test results obtained using the same test method by an infinite number of laboratories ($\sigma > 0$).

3.14
operator
person who normally and regularly carries out a particular test

3.15
outlier
result far enough in magnitude from other results to be considered not a part of the set

3.16
precision
closeness of agreement between the results obtained by applying the same test procedure several times on essentially the same materials and under prescribed conditions

Note 1 to entry: The smaller the random part of the experimental error, the more precise the procedure.

3.17
random error
component of measurement error that in replicate measurements varies in an unpredictable manner

3.18
repeatability
limiting value for the difference between two independent results obtained in the normal and correct operation of the same method, for test material considered to be the same, within a short interval of time, under the same test conditions, that is expected to be exceeded with a probability of 5% due to random variation

Note 1 to entry: Same test conditions are to be considered as same operator, same apparatus, same calibration and same laboratory.

Note 2 to entry: The representative parameter for the dispersion of the population that can be associated with these results is repeatability standard deviation or repeatability variance. Repeatability refers to the maximum difference attributable to random variation between two results obtained under the state of minimum random variability. Therefore, the period of time during which repeat results are to be obtained should be short enough to exclude time dependent variation, for example, variation caused by environmental changes, or variation associated with multiple calibrations".

Note 3 to entry: The term "repeatability" is not to be confused with the terms "between repeats" or "repeats".

3.19
reproducibility
limiting value for the difference between two independent results obtained in the normal and correct operation of the same method, for test material considered to be the same, under different test conditions, that is expected to be exceeded with a probability of 5 % due to random variation

Note 1 to entry: Different test conditions are to be considered as different operator, different apparatus, different calibration, and different laboratory.

Note 2 to entry: The representative parameter of the dispersion of the population that can be associated with these results is reproducibility standard deviation or reproducibility variance. Reproducibility refers to the maximum difference attributable to random variation between two results obtained under the state of maximum random variability.

3.20
result
final value obtained by following the complete set of instructions in a test method

Note 1 to entry: It is assumed that the result is rounded off according to the procedure specified in Annex G.

3.21
standard deviation
measure of the dispersion of a series of results around their mean, equal to the positive square root of the variance and estimated by the positive square root of the mean square

3.22
sum of squares
sum of squares of the differences between a series of results and their mean

3.23
true value
for practical purposes, the value towards which the average of single results obtained by *n* laboratories tends, as *n* tends towards infinity

Note 1 to entry: Such a true value is associated with the particular method of test.

Note 2 to entry: A different and idealized definition is given in ISO 3534-2[1].

3.24
variance
mean of the squares of the deviation of a random variable from its mean, estimated by the mean square

4 Stages in the planning of an interlaboratory study for the determination of the precision of a test method

4.1 General

The stages in planning an interlaboratory study (ILS) are as follows:

a) preparing a draft method of test;

b) planning a pilot study with at least two laboratories;

c) planning the ILS;

d) executing the ILS.

The four stages are described in turn in 4.2 to 4.5.

4.2 Preparing a draft method of test

This shall contain all the necessary details for carrying out the test and reporting the results. Any condition that could alter the results shall be specified.

The ILS shall be designed so that it covers the intended range of the test method (see also 6.5). A clause on precision is included in the draft method of the test at this stage only as a heading.

4.3 Planning a pilot study with at least two laboratories

A pilot study is necessary for the following reasons:

a) to verify the details in the operation of the test;

b) to find out how well operators can follow the instructions of the method, and thus of the ILS;

c) to check the precautions regarding samples;

d) to estimate approximately the precision of the test.

At least two samples are required, covering the range of results to which the test method is intended to apply; however, at least 12 laboratory/sample combinations shall be included. Each sample is tested twice by each laboratory under repeatability conditions. The samples should be equally distributed across the test method range, and should include major product groups covered in the test method scope. If any omissions or inaccuracies in the draft test method are revealed, they shall now be corrected. The results shall be analysed for precision, and bias for sample(s) with accepted reference values. If either is considered to be too large, then alterations to the test method shall be considered.

4.4 Planning the ILS

There shall be at least six participating laboratories, but it is recommended this number be increased to eight or more in order to ensure the final precision is based on at least six laboratories and to ensure the precision statement is more representative of the user population.

The number of samples shall be sufficient to adequately represent the types of materials to which the test method is to be applied, to cover the range of the property measured at approximately equidistant intervals, and to give reliability to the precision estimates. If precision is found to vary with the level of results in the pilot study, then at least five samples shall be used in the ILS. In order to correctly estimate precision versus level relationship, it is important that the choice of samples evenly covers the range and materials for the property measured, so that an estimated relationship is not too dependent upon the leverage of a sample with extreme property value.

It is strongly recommended that the leverage of each planned sample in the sample set design, lev_i, be assessed using Formula (2). No sample shall have a leverage exceeding 0,5. See Table D.11 for an example of leverage calculation (second column from the right under heading 'lev_i').

$$lev_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{k=1}^{n}(x_k - \bar{x})^2} \tag{2}$$

where

lev_i is leverage of sample i;

n is total number of planned samples;

x_i is Napierian logarithm, ln (p_i), with p_i being the planned property level for sample i;

$\bar{x}$ is grand average of all x_i.

In any event, it is necessary to obtain at least 30 degrees of freedom for both repeatability and reproducibility (see Annex B for the corresponding rationale). For repeatability, this means obtaining a total of at least 30 pairs of results in the ILS.

For reproducibility, Annex A, Table A.1 gives the minimum number of samples required in terms of L, P and Q, where L is the number of participating laboratories, and P and Q are the ratios of variance component estimates obtained from the pilot study. Specifically, P is the ratio of the interaction component to the repeats component and Q is the ratio of the laboratories component to the repeats component. Annex B gives the derivation of the formula used. If Q is much larger than P, then 30 degrees of freedom cannot be achieved; the blank entries in Table A.1 correspond to this situation (i.e. when more than 20 samples are required). For these cases, there is likely to be a significant bias between laboratories.

In the absence of pilot test program information to permit the use of Table A.1, the number of samples shall be greater than five, and chosen such that the number of laboratories times the number of samples is greater than or equal to 42.

When it is known or suspected that different types of materials exhibit different precision functional forms when tested by the test method, consideration should be given to conducting separate ILS for each type of material.

4.5 Executing the ILS

One person shall be responsible for the entire ILS, from the distribution of the texts of the test method and samples to the final appraisal of the results. This person shall be familiar with the test method, but shall not personally take part in the tests.

The text of the test method shall be distributed to all the laboratories in time to allow any queries to be raised before the tests begin. If any laboratory wants to practice the method in advance, than this shall be carried out with samples other than those used in the ILS.

The samples shall be accumulated, subdivided and distributed by the coordinator, who shall also keep a reserve of each sample for emergencies. It is most important that the individual laboratory portions be homogeneous and stable for the property of interest throughout the entire duration of the ILS. Prior to distribution, the ILS sample set shall be blind coded in a manner that preserves the anonymity of the nature of the test material and the expected value of the property. The following information shall be sent with the ILS sample set.

a) Agreed (draft) method of test.

b) Handling and storage requirements for the samples.

c) Order in which the samples are to be tested. A different random order for each laboratory is highly recommended. For large number of laboratories, several unique test orders may be randomly assigned to groups of laboratories, with no more than 4 laboratories per group.

d) For statistical reasons, it is imperative that the repeat results are obtained independently of each other, i.e. that the second result is not biased by knowledge of the first. This is achieved by blind coding where the repeat for each material in the ILS design is included in the test set sent to ILS participants without disclosing that it is a repeat, with an accompanying statement that a single result is to be obtained on each sample in the test set, in the specified testing order, by the same operator with the same apparatus within a short time. If this blind coding is regarded as infeasible to achieve, then the statement shall state that a pair of results associated with a sample shall be obtained by the same operator with the same apparatus within a short time, without disclosing the nature of the sample.

e) Period of time within which all the samples are to be tested.

f) Blank form for reporting the results. For each sample, there shall be space for the date of testing, the test results, and any unusual occurrences. The unit of accuracy for reporting the results shall be specified.

g) Statement that the test shall be carried out under normal conditions, using qualified operators who carry out this kind of test routinely and that the duration of the test shall be the same as normal.

h) A questionnaire requesting information on the conditions used in the application of the test method, e.g. apparatus details, reagents and materials, calibration and verification procedures, quality control procedure, any deviations from either the test method or the instructions supplied, observations and suggestions for future improvement of the test method.

Operators that participated in the pilot study may also participate in the ILS. If their extra experience in testing a few more samples produces a noticeable effect, it serves as a warning that the test method is not satisfactory. They shall be identified in the report of the results so that any effect can be noted.

NOTE For additional guidance on the planning and execution of an ILS, consult ASTM D7778[4] and ASTM D6300[2].

5 Statistical treatment of ILS results

5.1 General recommendation

Although the procedures described in Clauses 5 and 6 of this document are in a form suitable for hand calculation, it is strongly advised that these procedures be carried out using an electronic computer with appropriately validated software designed specifically to store and analyse ILS test results based on the procedures of this document. It is also highly recommended that these procedures be carried out under the guidance of a statistician.

NOTE A software package extensively used in the ISO and ASTM community is D2PP[13]. That software package does not include GESD or Cook's Distance assessment in line with this document.

In the clauses to follow, procedures are specified to achieve the following:

a) pre-screen the results as reported from the ILS on a sample-by-sample basis for grossly discordant results (outliers);

b) assess independence or dependence of precision and the level of results after pre-screening;

c) assess uniformity of precision from laboratory to laboratory by detecting the presence (or absence) of additional outliers using the detection power from the entire data set.

The procedures are described in mathematical terms based on the notation of Annex C.

Illustration of the procedures is provided in referenced Annexes.

For all the procedures, it is assumed that the results are either from a single normal distribution or capable of being transformed into such a distribution (see 5.3). Other cases (which are rare) require a different treatment that is beyond the scope of this document. See Reference [6] for a statistical test on normality.

5.2 Pre-screen using GESD technique

Prior to execution of 5.3 to 5.7, examine all information returned by ILS participants to determine compliance with agreed-upon test protocol and method of test. If the investigation disclosed no clerical, sampling or procedural errors, apply the Generalized Extreme Studentized Deviation (GESD) technique as outlined in this clause to results received for each ILS sample to identify unusual or extreme results. Investigation for causes associated with unusual results shall be conducted. If acceptable cause(s) is found during the investigation, the unusual results shall be either corrected, replaced, or rejected. Correction or replacement of the unusual results with a new set of results shall be approved by the ILS

coordinator in consultation with the ILS statistician. If no acceptable cause is found, the unusual or extreme results as identified by the GESD technique at the 99 % confidence level shall be rejected.

An overall summary of this GESD pre-screening technique is outlined below.

For each ILS sample, execute the following steps.

1) Calculate the sample mean using all results received for the sample.

2) Calculate difference for each pair of results as received from laboratories that have reported both results.

3) Identify outlier(s) in the data set of differences obtained from step 2) by following the methodology outlined in Annex H.

4) For each outlying difference identified, remove the member from the pair that is farthest from the sample mean calculated in 1) and replace it with the value of the remaining result.

5) For laboratories that have only reported one result, i.e. the other result is missing, assign the value of the single reported result to the missing result before proceeding to step 6).

6) Calculate the sum of the pair of the results for each lab. For laboratories that have reported both results and neither result has been rejected, this will be the sum of both reported results. In the case where one of the pair of results is missing (not reported) or rejected from step 4), this sum will be twice the single reported result since the missing result is assigned the same value as the reported result.

7) Identify outlier(s) in data set of sums as obtained from step 6) by following the methodology outlined in Annex H.

8) For each outlying sum of results, exclude both results from further statistical analysis.

9) For the pairs of results with sums that have not been rejected, retain both reported results for analysis if both results are as originally received from the laboratories. If one of the two results of the pair is an assigned value from step 4) or step 5), retain the reported result from the laboratories for analysis, and treat the other result as "missing".

10) The data set remaining after completion of step 9) then constitutes the data set to be further analysed as per 5.3 to 5.7.

5.3 Transformation of data and outlier tests

5.3.1 General

In many test methods, the precision depends on the level of the test result, and thus the variability of the reported results is different from sample to sample. The method of analysis outlined in this document requires that this shall not be so and the position is rectified, if necessary, by a transformation.

The laboratories standard deviations, D_j, and the repeats standard deviations, d_j, for sample j (see Annex C for notation explanation) are calculated and plotted separately against the sample means, m_j, in accordance with Annexes D and E). If the points so plotted can be considered as lying about a pair of lines parallel to the m-axis, then no transformation is necessary. If, however, the plotted points describe non-horizontal straight lines or curves of the form $D = f_1(m)$ and $d = f_2(m)$, then a transformation is necessary.

The relationships $D = f_1(m)$ and $d = f_2(m)$ are not, in general, identical. The statistical procedures of this document require, however, that the same transformation be applicable both for repeatability and for reproducibility. For this reason, the two relationships are combined into a single dependency relationship $D = f(m)$ (where D now includes d) by including a dummy variable, T. This takes account of the difference between the relationships, if one exists, and provides a means of testing for this difference (see F.1).

The single relationship $D = f(m)$ is best estimated by a weighted linear regression analysis, even though in most cases an unweighted regression gives a satisfactory approximation. The derivation of weights is described in F.2, and the computational procedure for the regression analysis is described in F.3. Typical forms of dependence $D = f(m)$ are given in E.1. These are all expressed in terms of transformation parameters B and B_0.

The estimation of B and B_0, and the transformation procedure which follows, are summarized in E.2. This includes statistical tests for the significance of the regression (i.e. is the relationship $D = f(m)$ parallel to the m-axis), and for the difference between the repeatability and reproducibility relationships, based at the 5 % significance level. If such a difference is found to exist, or if no suitable common transformation exists, then the alternative sample by sample procedures of ISO 5725-2 shall be used. In such an event, it is not possible to test for laboratory bias over all samples (see 5.6) or separately estimate the interaction component of variance (see 6.2).

If it has been shown at the 5 % significance level that there is a significant regression of the form $D = f(m)$, then the appropriate transformation $y = F(x)$, where x is the reported result, is given by Formula (3):

$$F(x) = K \int \frac{dx}{f(x)} \tag{3}$$

where K is a constant.

In that event, all results shall be transformed accordingly and the remainder of the analysis carried out in terms of the transformed results. Typical transformations are given in E.1.

It is difficult to make the choice of transformation the subject of formalized rules. Qualified statistical assistance can be required in particular cases. The presence of outliers can affect judgement as to the type of transformation required, if any (see 5.7). That is why extremely discordant results shall be removed as described in 5.1 above prior to making a judgement on transformation(s).

The transformation and outlier procedure is described in the form of a flow chart in Figure 1. Note that the transformation process is an iterative procedure, requiring confirmation of the choice of transformation if outliers have been rejected. If the original transformation is found to be inadequate after outliers have been removed, then a different transformation will be required.

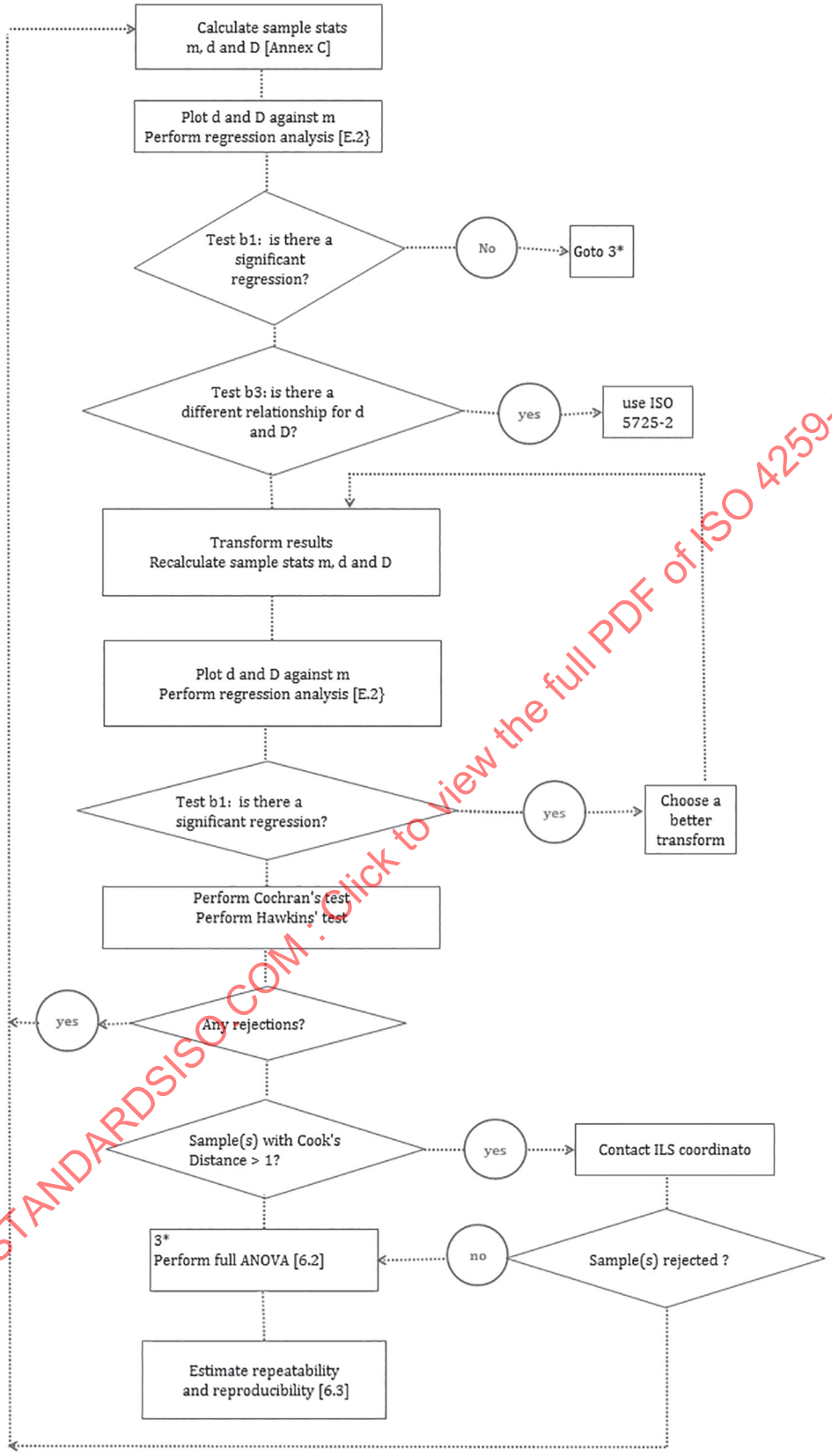


Figure 1 — Transformation and outlier procedure

5.3.2 Outlier identification after pre-screening

The pre-screened results, or if it has been decided that a transformation is necessary, the pre-screened and transformed results, shall be further tested statistically for outliers. These are the values that are so different from the remaining data that it can only be concluded that they have arisen from some fault in the application of the test method or from testing a wrong sample. Many possible tests may be used and the associated significance levels can be varied, but those that are given below have been found to be appropriate for this document. These outlier tests all assume a normal distribution of errors (see 5.1).

5.3.3 Uniformity of repeatability

The first outlier test is concerned with detecting a discordant result for the absolute difference between a pair of repeat results. This test[7] involves calculating the e^2_{ij} over all the laboratory/sample combinations. Cochran's criterion at the 1 % significance level is then used to test the ratio of the largest of these e^2_{ij} values over their sum (see C.5). If its value exceeds the value given in Table D.14, corresponding to one degree of freedom, n_r being the number of pairs available for comparison, then the member of the pair farthest from the sample mean shall be rejected and the process repeated, reducing n_r by 1, until no more rejections are called for. In certain cases, this test "snowballs" and leads to an unacceptably large proportion of rejections (say more than 10 %). If this is so, this rejection test shall be abandoned and some or all of the rejected results shall be retained. An arbitrary decision based on judgement is necessary in this case. See D.3 for an illustration.

5.3.4 Uniformity of reproducibility

The following outlier test, Hawkins' test, see also D.4, is concerned with establishing uniformity in the reproducibility estimate. It is designed to detect a discordant pair of results from a laboratory on a particular sample. It involves, for each sample, forming the ratio of the largest absolute deviation of laboratory mean from the overall sample mean to the square root of certain sums of squares (see C.6).

The ratio corresponding to the largest absolute deviation shall be compared with the critical 1 % values given in Table D.15, where n_R is the number of laboratory cells in the sample concerned and where ν is the degrees of freedom for the sum of squares, which is additional to that corresponding to the sample in question. That is, for this test, ν_R refers to the degrees of freedom from other samples (i.e. excludes the sample being tested).

If a significant value is encountered, the corresponding extreme value shall be omitted and the process repeated.

If the test "snowballs", leading to an unacceptably large proportion of rejections (say more than 10 %), then this rejection test shall be abandoned and some or all of the rejected results shall be retained. An arbitrary decision based on judgement is necessary in this case.

5.4 Rejection of complete data (from all laboratories) for a sample

The laboratories standard deviation and repeats standard deviation shall be examined for any outlying samples. If a transformation has been carried out or any rejection made, new standard deviations shall be calculated. See D.5 for further illustration.

If the standard deviation for any sample is excessively large, it shall be examined with a view to rejecting the results from that sample.

Cochran's criterion at the 1 % level can be used when the standard deviations are based on the same number of degrees of freedom. This involves calculating the ratio of the largest of the corresponding sums of squares (laboratories or repeats, as appropriate) to their total (see C.5). If the ratio exceeds the critical value given in Table D.14, with n as the number of samples and ν the degrees of freedom, then all the results from the sample in question shall be rejected. In such an event, care should be taken that the extreme standard deviation is not due to the application of an inappropriate transformation (see 5.3), or undetected outliers.

There is no optimal test when standard deviations are based on different degrees of freedom. However, the variance ratio (i.e. the ratio of largest variance to that pooled from the remaining samples) follows an *F*-distribution with v_1 and v_2 degrees of freedom (see C.7). Here v_1 is the degrees of freedom of the variance in question and v_2 is the degrees of freedom for the remaining samples. If the ratio is greater than the critical value given in Tables D.17 to D.20, corresponding to a significance level of 0,01/S, where S is the number of samples, then results from the sample in question shall be rejected.

5.5 Estimating missing or rejected values

5.5.1 One of the two repeat values missing or rejected

If one of a pair of repeats (x_{ij1} or x_{ij2} for un-transformed results, y_{ij1} or y_{ij2} for transformed results) is missing or rejected, this shall be considered to have the same value as the other repeat in accordance with the least squares method.

5.5.2 Both repeat values missing or rejected

If both the repeat values are missing, estimates of a_{ij} [(x_{ij1} + x_{ij2}) for un-transformed results, or (y_{ij1} + y_{ij2}) for transformed results] shall be made by forming the laboratories × samples interaction sum of squares, including the missing values of the totals of the laboratories/samples pairs of results as unknown variables. Any laboratory or sample from which all the results were rejected shall be ignored and new values of L and S used. The estimates of the missing or rejected values shall then be found by forming the partial derivatives of this sum of squares with respect to each variable in turn and equating these to zero to solve as a set of simultaneous formulae.

Formula (4) may be used where only one pair sum has to be estimated. If more estimates are to be made, the technique of successive approximation can be used. In this, each pair sum is estimated in turn from Formula (4), using L_1, S_1 and T_1 values which contain the latest estimates of the other missing pairs. Initial values for estimates can be based on the appropriate sample mean, and the process usually converges to the required level of accuracy within three complete iterations. See Reference [9] for details.

If the value of one pair sum, a_{ij}, has to be estimated, the estimate is given by Formula (4):

$$a_{ij} = \frac{1}{(L-1)(S'-1)} \times (LL_1 + S'S_1 - T_1) \tag{4}$$

where

S' is S minus the number of samples rejected in 5.4;

L_1 is the total of remaining pairs in the ith laboratory;

S_1 is the total of remaining pairs in the jth sample;

T_1 is the total of all pairs except a_{ij}.

See illustration in D.6 for estimating a single pair of missing values.

5.6 Rejection test for outlying laboratories

At this stage, one further rejection test remains to be carried out. This determines whether it is necessary to reject the complete set of results from any particular laboratory (i.e. a discordant set of results from a laboratory on all samples). It cannot be carried out at an earlier stage, except in the case where no individual results or pairs are missing or rejected. The procedure again consists of Hawkins' test (see 5.3.4), applied to the laboratory averages over all samples. If any laboratories are rejected on all samples, new estimates shall be calculated for any remaining missing values (see 5.5).

This test involves identifying and forming the ratio of the largest absolute deviation of laboratory-average-over-all samples versus the overall mean to the square root of certain sums of squares (see C.6).

For this test, n is the total number of laboratories, v is zero. See illustration in D.7.

5.7 Confirmation of selected transformation

5.7.1 General

At this stage, it is necessary to check that the rejections carried out have not invalidated the transformation used. If necessary, the procedure given in 5.3 shall be repeated with the outliers deleted, and if a new transformation is selected, outlier tests shall be reapplied. See also D.8.

5.7.2 Identification of excessively influential sample(s)

The last step prior to proceeding with analysis of variance and calculation of precision estimates in Clause 6 is to determine if the selection of transformation function is excessively influenced by one or more samples. Cook's Distance is the recommended statistic for this evaluation. Cook's Distance is calculated for every sample in the unweighted linear regression of $\ln(D_i)$ versus $\ln(m_i)$ using the untransformed ILS results excluding the outliers identified in 5.2 to 5.6. Cook's Distance is a metric which combines the leverage (lev_i, see 4.4) of a sample along with the degree of fit with and without use of this sample in the regression. This will determine if the regression relationship is overly dependent on the sample. A sample with Cook's Distance exceeding 1 constitutes a highly influential sample, and is a candidate for exclusion. The ILS coordinator shall be notified of any sample with Cook's Distance exceeding 1. Exclusion of samples based on Cook's Distance shall be discussed with the ILS coordinator, who shall make the final decision after consultation with all stakeholders and the statistician.

Cook's Distance is calculated as follows:

$$\text{Cook's Distance} = \frac{r_i^2}{p} \times \frac{lev_i}{(1 - lev_i)} \tag{5}$$

where

p = 2 (for regression with slope and intercept);

lev_i is leverage of sample i [see Formula (2)];

r_i is studentized residual of sample i [see Formula (6)].

$$r_i = \frac{\varepsilon_i}{s(i)\sqrt{1 - lev_i}} \tag{6}$$

where

ε_i is the residual of sample i;

$s(i)$ is the residual mean square obtained from regression with the exclusion of sample i.

$s(i)$ can be calculated by solving Formula (7).

$$(n-3)[s(i)]^2 = (n-2)s^2 - \frac{\varepsilon_i^2}{(1 - lev_i)} \tag{7}$$

where

$$s^2 = \sum_{i=1}^{n} \varepsilon_i^2 \Big/ (n-2)$$

6 Analysis of variance, calculation and expression of precision estimates

6.1 General

After the data have been inspected for uniformity, a transformation has been performed if necessary, and any outliers have been rejected (see Clause 5), an analysis shall be carried out. First an analysis of variance table shall be constructed, and finally the precision estimates derived.

6.2 Analysis of variance

6.2.1 Forming the sums of squares for the laboratories × samples interaction sum of squares

The estimated values, if any, shall be put in the array and an approximate analysis of variance performed.

Mean correction, $M_c = (TOT)^2/2L'S'$ (8)

where L' is L minus the number of laboratories rejected in 5.6 minus the number of laboratories with no remaining results after rejections in 5.3.4.

$$\text{Samples sum of squares} = \left[\sum_{j=1}^{S'} (g_j^2/2L')\right] - M_c \quad (9)$$

$$\text{Laboratories sum of squares} = \left[\sum_{i=1}^{L'} (h_i^2/2S')\right] - M_c \quad (10)$$

$$\text{Pairs sum of squares} = (1/2)\left[\sum_{i=1}^{L'}\sum_{j=1}^{S'} a_{ij}^2\right] - M_c \quad (11)$$

The laboratories × samples interaction sum of squares, I, is given by:

I = (pairs sum of squares) − (laboratories sum of squares) − (sample sum of squares)

Ignoring any pairs in which there are estimated values,

$$E = \text{repeats sum of squares} = (1/2)\sum_{i=1}^{L'}\sum_{j=1}^{S'} e_{ij}^2 \quad (12)$$

The purpose of performing this approximate analysis of variance is to obtain the minimized laboratories × samples interaction sum of squares, I. This is then used as indicated in 6.2.2, to obtain the laboratories sum of squares. See D.9 for further guidance.

If there were no estimated values, the above analysis of variance is exact and 6.2.2 shall be disregarded.

6.2.2 Forming the sum of squares for the exact analysis of variance

In this sub-clause, all the estimated pairs are disregarded and new values of g_j are calculated. The following sums of squares for the exact analysis of variance[10] are formed. See further D.10.

$$\text{Uncorrected sample sum of squares} = \sum_{j=1}^{S'} \frac{g_j^2}{S_j} \tag{13}$$

where S_j is 2 (L' – number of missing pairs in that sample).

$$\text{Uncorrected pairs sum of squares} = (1/2)\sum_{i=1}^{L'}\sum_{j=1}^{S'} a_{ij}^2 \tag{14}$$

The laboratories sum of squares is equal to (pairs sum of squares) – (samples sum of squares) – (the minimized laboratories × samples interaction sum of squares).

$$= (1/2)\left[\sum_{i=1}^{L'}\sum_{j=1}^{S'} a_{ij}^2\right] - \left[\sum_{j=1}^{S'} \frac{g_j^2}{S_j}\right] - I \tag{15}$$

6.2.3 Degrees of freedom

The degrees of freedom for the laboratories are (L' – 1). The degrees of freedom for laboratories × samples interaction are (L' – 1) (S' – 1) for a complete array and are reduced by one for each pair which is estimated. The degrees of freedom for repeats are ($L'S'$) and are reduced by one for each pair in which one or both values are estimated. See also D.11.

6.2.4 Mean squares and analysis of variance

The mean square in each case is the sum of squares divided by the degrees of freedom. This leads to the analysis of variance shown in the following Table 1.

Table 1 — Analysis of variance

Source of variation	Degrees of freedom	Sum of squares	Mean square
Laboratories	$L' - 1$	Laboratories sum of squares	M_L
Laboratories × samples	$(L' - 1)(S' - 1)$ – number of estimated pairs	I	M_{LS}
Repeats	$L'S'$ – number of pairs in which one or both values are estimated	E	M_r

The ratio M_L/M_{LS} is distributed as F with the corresponding laboratories and interaction degrees of freedom (see C.7). If this ratio exceeds the 5 % critical value given in Table D.17, then bias between the laboratories is implied and the ILS coordinator shall be informed (see 4.5): further standardization of the test method can be necessary. See also D.12.

6.3 Expectation of mean squares and calculation of precision estimates

6.3.1 Expectation of mean squares with no estimated values

For a complete array with no estimated values, the expectations of mean squares are:

Laboratories (M_L): $\sigma_0^2 + 2\sigma_1^2 + 2S'\sigma_2^2$

Laboratories × samples (M_{LS}): $\sigma_0^2 + 2\sigma_1^2$

Repeats (M_r): σ_0^2

where

σ_0^2 is the component of variance due to repeats;

σ_1^2 is the component of variance due to interaction between laboratories and samples;

σ_2^2 is the component of variance due to differences between laboratories.

6.3.2 Expectation of mean squares with estimated values

The coefficients of σ_0^2 and σ_2^2 in the expectation of mean squares are altered in the cases where there are estimated values. The expectations of mean squares then become:

Laboratories (M_L): $\alpha\sigma_0^2 + 2\sigma_1^2 + \beta\sigma_2^2$

Laboratories × samples (M_{LS}): $\gamma\sigma_0^2 + 2\sigma_1^2$

Repeats (M_r): σ_0^2

where

$I2 = 2\dfrac{(L_N - S')}{L'-1}$ $\beta = 2\dfrac{(K - S')}{(L'-1)}$

K is the number of laboratory × sample cells containing at least one result;

α and γ are computed as follows.

a) If there are no cells with only a single estimated result, then $\alpha = \gamma = 1$.

b) If there are no empty cells (i.e. every laboratory has tested every sample at least once, and $K = L' \times S'$), then α and γ are both 1 plus the proportion of cells with only a single result.

c) If there are both empty cells and cells with only one result, then for each laboratory compute the proportion, p_i, of samples tested for which there is only one result, and the sum, P, of these proportions over all laboratories. For each sample, compute the proportion, q_j, of laboratories that have tested the sample for which there is only one result, and the sum, Q, of these proportions over all samples. Compute the total number of cells, W, with only one result and the proportion of these among all non-empty cells, W/K. Then:

$$\alpha = 1 + \frac{P - (W/K)}{L'-1} \tag{16}$$

and

$$\gamma = 1 + \frac{W - P - Q + (W/K)}{K - L' - S' + 1} \tag{17}$$

NOTE The above mentioned development is based upon the assumption that both samples and laboratories are "random effects".

See illustration in Annex D.

6.3.3 Calculation of precision estimates

6.3.3.1 Repeatability

The repeatability variance, V_r, is twice the mean square for repeats. The repeatability estimate is the product of the repeatability standard deviation and the "t-value", t_v, with appropriate degrees of freedom, v_r (see Table D.16), corresponding to a two sided probability of 95 %.

This calculated estimate shall be rounded to no fewer than three and no more than four significant digits.

Note that if a transformation $Y = F(x)$ has been used, then

$$r(x) = \left|\frac{dx}{dy}\right| r(y) \tag{18}$$

where $r(x)$, $r(y)$ are the corresponding repeatability functions for untransformed (x) and transformed (y) results in accordance with Table E.1.

A similar relationship applies to the reproducibility functions $R(x)$, $R(y)$.

6.3.3.2 Reproducibility

The reproducibility variance, V_R, is expressed as

$$V_R = 2(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)$$

and can be calculated using Formula (19):

$$V_R = \frac{2}{\beta} M_L + \left(1 - \frac{2}{\beta}\right) M_{LS} + \left(2 - \gamma + \frac{2}{\beta}(\gamma - \alpha)\right) M_r \tag{19}$$

where the symbols are as set out in 6.2.4 and 6.3.3.

The reproducibility estimate is the product of the reproducibility standard deviation and the "t-value", t_v, with appropriate degrees of freedom, v_R, (see Table D.16), corresponding to a two sided probability of 95 %. An approximation[11] to the degrees of freedom of the reproducibility variance, v_R, is given by Formula (20).

$$v_R = \frac{V_R^2}{\frac{r_1^2}{L' - 1} + \frac{r_2^2}{v_{LS}} + \frac{r_3^2}{v_r}} \tag{20}$$

where

v_{LS} is the degrees of freedom for laboratories × samples;

v_r is the degrees of freedom for repeats.

r_1, r_2 and r_3 are the three successive terms in Formula (18), i.e:

$$r_1 = \frac{2}{\beta} ML$$

$$r_2 = \left(1 - \frac{2}{\beta}\right) M_{LS}$$

and

$$r_3 = \left(2 - \gamma + \frac{2}{\beta}(\gamma - \alpha)\right) M_r$$

The calculated estimate of reproducibility shall also be rounded to no fewer than three and no more than four significant digits. See D.14 for further details.

Substantial bias between laboratories results in a loss of degrees of freedom estimated by Formula (20). If reproducibility degrees of freedom are less than 30, then the ILS coordinator shall be informed (see 4.5); further standardization of the test method can be necessary.

It is a statistical expectation that repeatability is less than reproducibility. In the rare event where the numerical value of reproducibility is less than the repeatability value, the statistical evaluation should be verified. If no discrepancies are found, the numerical value for repeatability shall also be used for reproducibility. This also implies that the test method under examination shows serious deficits and needs to be checked carefully and/or revised. For this special situation, the method should not be used as a specification method.

6.4 Expression of precision estimates of a method of test

6.4.1 When the precision of a method of test has been determined in accordance with the procedures set out in this document, it shall be included in the method as outlined below. The range (x to y) in X.1 below, shall be determined based on lowest and highest achieved non-rejected results from the ILS.

X Precision

X.1 General

The precision, as determined by statistical examination in accordance with ISO 4259 of interlaboratory test results on (type of products) with test results in the range (x to y), is given in X.2 and X.3.

X.2 Repeatability

The difference between two independent results obtained in the normal and correct operation of the same method, for test material considered to be the same, within a short interval of time, under the same test conditions, that is expected to be exceeded with a probability of 5 % due to random variation, can be calculated using the following function:

$$r = f_r(x)$$

where x is the average of the two test results being compared.

X.3 Reproducibility

The difference between two independent results obtained in the normal and correct operation of the same method, for test material considered to be the same, under different test conditions, that is expected to be exceeded with a probability of 5 % due to random variation, can be calculated using the following function:

$$R = f_R(x)$$

where x is the average of the two test results being compared.

6.4.2 Only in exceptional cases shall a precision estimate not based upon or not meeting ISO 4259 requirements be allowed. In those cases, the alternative introductory text below shall be used:

"The precision evaluation study for the matrix of samples with (*p*) contents in the range (*q* to *r*) did not conform to the requirements of ISO 4259, and thus only an estimate of precision based upon interlaboratory test results is given in X.2 and X.3."

(*p* is the property name; *q* and *r* are the lowest and highest achieved result in the ILS for *p*.)

The rationale for accepting the precision estimate, along with the nature of the noncompliance (if any), shall accompany the above statement.

6.4.3 A summary description of the ILS used to generate the precision statement shall be provided. At minimum, the summary description shall include the number of laboratories, number and type of materials studied, and range of the measured average property levels. The summary may be included as a Note.

6.5 Specification of scope for the test method

The scope limits for the test method shall be specified based on the *r* and *R* achieved from the ILS for the draft test method, the measurement range, and the reporting resolution (in accordance with Annex G) as follows.

The lower limit of the scope of the test method shall be the larger of lowest sample mean tested in the ILS or lowest achievable result + 2**R*, where *R* is evaluated at the lowest sample mean. The lowest achievable result is defined as the point of truncation (excluding minus infinity) below which a result is not defined. An example is the zero point for a percent concentration. Due to testing variation, the lowest acceptable single result that is deemed as a valid result of the test method shall be lower method scope limit – 1,2**R*, where *R* is evaluated at the low method scope limit value.

Similarly, the upper limit of the scope of a test method shall be the lesser of highest sample mean tested in the ILS or highest achievable result – 2*R*, where *R* is evaluated at the highest sample mean. The highest achievable result is defined as the point of truncation (excluding plus infinity) above which a result is not defined. An example is the 100 point for a percent concentration. Due to testing variation, the highest acceptable single result that is deemed as a valid result of the test method shall be higher method scope limit + 1,2**R*, where *R* is evaluated at the high method scope limit value.

The matter is clarified in Figure 2.

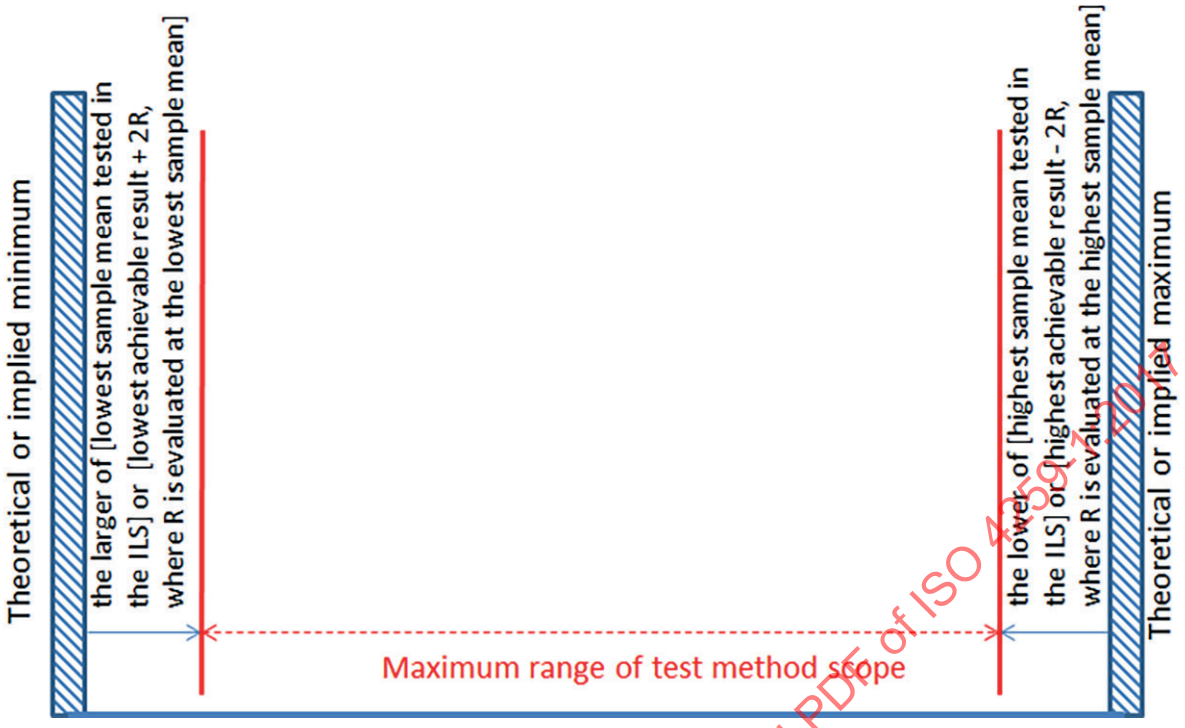


Figure 2 — Scope limit setting clarified

7 *R/r* ratio

The *R/r* ratio can be used as an approximate indicator of the contribution of bias between laboratories towards reproducibility, *R*. A large *R/r* ratio is an indication that the lab-bias effect is a significantly dominant contributor towards *R* over the random effects. When the lab-bias effect is the significantly dominant contributor for *R*, it means that a large portion of the difference in results between laboratories are not due to random testing noise, but rather, due to systemic bias between laboratories. When lab-bias effect is significantly dominant, this suggests further between-lab standardization protocol is necessary to reduce *R*. If the *R/r* ratio exceeds 4, this situation shall be brought to the attention of test method developers since this constitutes compelling statistical evidence that further between lab standardization protocol is necessary.

NOTE An *R/r* ratio > 4 criterion is based on extensive experience in performing of analyses of petroleum industry test methods in line with this document.

Annex A
(normative)

Determination of number of samples required

This annex describes the number of samples required for the planning of an ILS as described in 4.4.

Table A.1 — Determination of number of samples required to meet 30 degrees of freedom

L = 6

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	3									
1	4	11								
2	5	7								
3	5	7	14							
4	5	6	10							
5	6	6	8	15						
6	6	6	8	11						
7	6	6	7	10	15					
8	6	6	7	9	12					
9	6	6	7	8	10	15				

L = 7

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	3									
1	4	7								
2	4	6	17							
3	4	5	9							
4	5	5	7	13						
5	5	5	6	9	19					
6	5	5	6	8	12					
7	5	5	6	7	10	15				
8	5	5	6	7	8	12	20			
9	5	5	6	6	8	10	14			

L = 8

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	3									
1	3	5								
2	4	5	9							
3	4	5	7	14						
4	4	4	6	9	20					
5	4	4	5	7	11					
6	4	4	5	6	8	13				
7	4	4	5	6	7	10	16			
8	4	5	5	6	6	8	11	18		
9	4	5	5	5	6	7	9	13		

L = 9

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	2									
1	3	4								
2	3	4	7							
3	3	4	5	9						
4	4	4	5	6	11					
5	4	4	5	6	7	12				
6	4	4	4	5	6	9	14			
7	4	4	4	5	6	7	10	15		
8	4	4	4	5	5	6	8	10	16	
9	4	4	4	5	5	6	7	8	11	18

L = 10

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	2	8								
1	3	4	11							
2	3	4	5	12						
3	3	3	4	6	13					
4	3	4	4	5	7	14				
5	3	4	4	5	6	8	14			
6	3	4	4	4	5	6	9	14		
7	3	4	4	4	5	6	7	9	14	
8	3	4	4	4	5	5	6	7	10	14
9	4	4	4	4	4	5	6	6	8	10

L = 11

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	2	4								
1	2	3	5							
2	3	3	3	7						
3	3	3	4	5	8					
4	3	3	4	4	6	8	18			
5	3	3	4	4	5	6	9	15		
6	3	3	3	4	4	5	6	9	14	
7	3	3	3	4	4	5	5	7	9	13
8	3	3	3	4	4	4	5	6	7	9
9	3	3	3	4	4	4	5	5	6	7

L = 12

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	2	4								
1	2	3	5							
2	2	3	4	6	14					
3	3	3	3	4	6	11				
4	3	3	3	4	5	6	9			
5	3	3	3	4	4	5	6	9	16	
6	3	3	3	3	4	4	5	6	9	13
7	3	3	3	3	4	4	5	5	6	8
8	3	3	3	3	4	4	4	5	5	6
9	3	3	3	3	3	4	4	4	5	6

L = 13

P: \ *Q*:	0	1	2	3	4	5	6	7	8	9
0	2	3								
1	2	3	4	12						
2	2	3	3	4	8					
3	2	3	3	4	5	7	14			
4	3	3	3	3	4	5	7	10		
5	3	3	3	3	4	4	5	6	9	15
6	3	3	3	3	3	4	4	5	6	8
7	3	3	3	3	3	4	4	4	5	6
8	3	3	3	3	3	3	4	4	5	5
9	3	3	3	3	3	3	4	4	4	5

Table A.1 *(continued)*

$L = 14$										
	Q:	0	1	2	3	4	5	6	7	8	9
P:	0	2	3								
	1	2	2	3	7						
	2	2	2	3	4	6	12				
	3	2	2	3	3	4	5	8	18		
	4	2	3	3	3	3	4	5	7	11	
	5	2	3	3	3	3	4	4	5	6	8
	6	3	3	3	3	3	3	4	4	5	6
	7	3	3	3	3	3	3	3	4	4	5
	8	3	3	3	3	3	3	3	4	4	4
	9	3	3	3	3	3	3	3	3	4	4

$L = 15$										
	Q:	0	1	2	3	4	5	6	7	8	9
P:	0	2	2	13							
	1	2	2	3	5	19					
	2	2	2	3	3	4	7				
	3	2	2	3	3	3	4	6	9		
	4	2	2	3	3	3	4	4	5	7	10
	5	2	2	3	3	3	3	4	4	5	6
	6	2	2	3	3	3	3	3	4	4	5
	7	2	2	3	3	3	3	3	3	4	4
	8	2	2	3	3	3	3	3	3	3	4
	9	2	2	3	3	3	3	3	3	3	3

$L = 16$										
	Q:	0	1	2	3	4	5	6	7	8	9
P:	0	1	2	5							
	1	2	2	3	4	8					
	2	2	2	2	3	4	5	9			
	3	2	2	2	3	3	4	4	6	9	
	4	2	2	2	3	3	3	4	4	5	6
	5	2	2	2	3	3	3	3	4	4	5
	6	2	2	2	3	3	3	3	3	4	4
	7	2	2	2	3	3	3	3	3	3	4
	8	2	2	2	3	3	3	3	3	3	3
	9	2	2	2	3	3	3	3	3	3	3

L = number of participating laboratories

$$P = \frac{\text{interaction variance component}}{\text{repeats variance component}}$$

$$Q = \frac{\text{laboratories variance component}}{\text{repeats variance component}}$$

Annex B
(informative)

Derivation of formula for estimating the number of laboratories and samples required to meet minimum 30 degrees of freedom

B.1 Degrees of freedom

This annex explains the rationale behind the necessary 30 degrees of freedom (Table A.1) as required by 4.4.

An analysis of variance is carried out on the results of the pilot study. This yields rough estimates of the three components of variance due to interaction between laboratories and samples, namely:

— σ_0^2 for repeats;

— σ_1^2 for laboratories × samples interaction;

— σ_2^2 for laboratories.

Substituting the above into Formula (20) (see 6.3.3.2) for calculating the reproducibility degrees of freedom, this becomes:

$$\frac{(1+P+Q)^2}{v} = \frac{\left[\left(\frac{1}{2}+P\right)\Big/S+Q\right]^2}{(L-1)} + \frac{(S-1)\left(\frac{1}{2}+P\right)^2}{S^2(L-1)} + \frac{1}{4LS} \tag{B.1}$$

where

P is the ratio σ_1^2/σ_0^2;

Q is the ratio σ_2^2/σ_0^2;

v is the degrees of freedom of the reproducibility variance;

L is the number of laboratories;

S is the number of samples.

The formula rearranges into the form:

$aS + b = 0$

where

$a = vQ^2 - (1 + P + Q)^2(L - 1)$

$$b = v\left[\left(2Q + \frac{1}{2} + P\right)\left(\frac{1}{2} + P\right) + 0{,}25(L-1)/L\right]$$

Therefore

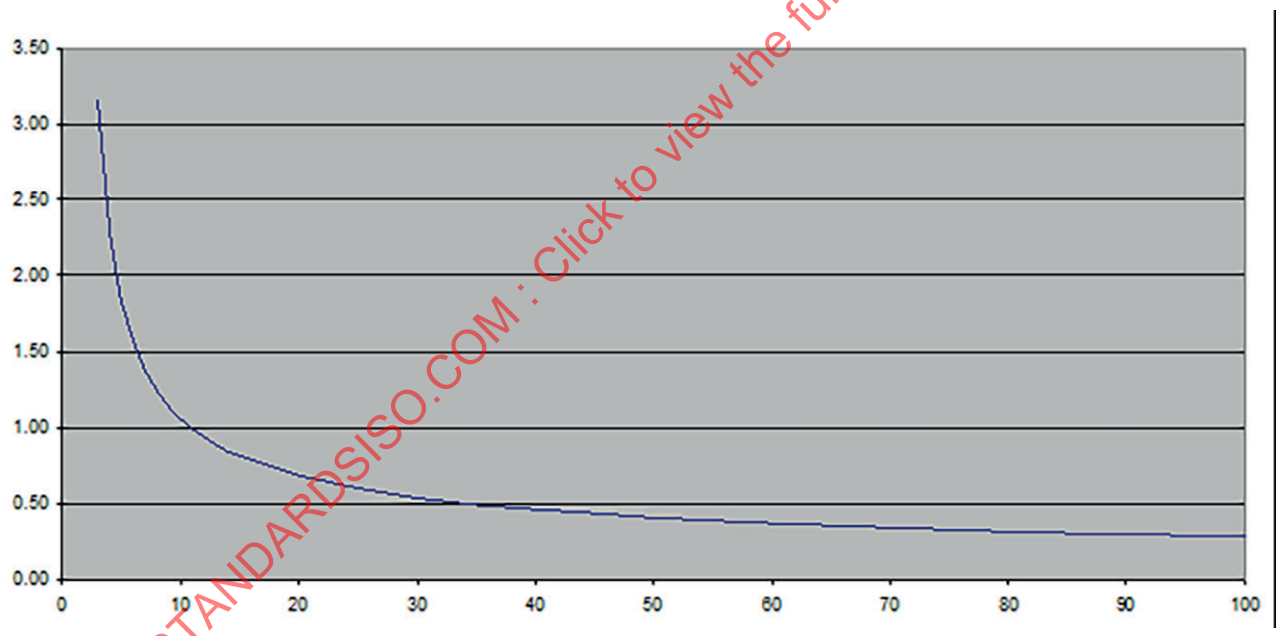
$$S = -\frac{b}{a} \tag{B.2}$$

gives the values of S for given values of L, P, Q and v.

Table A.1 is based on v = 30 degrees of freedom. For non integral values of P and Q, S can be estimated by second order interpolation from Table A.1.

B.2 Explanation for choice of 30 as minimum degrees of freedom

The above procedure is based upon v = 30 degrees of freedom. The reason why the number 30 is chosen is explained in Figure B.1, which plots the ratio of the 95 % confidence interval width of a standard deviation estimate for σ^{-1} against degrees of freedom. Because precision estimates are calculated using standard deviation estimates, they will have almost identical confidence interval relationships. Given that a confidence interval quantifies the level of uncertainty which we have in an estimate, it is clear that beyond 30 degrees of freedom the rate of improvement (reduction in confidence interval width) in this respect is minimal. Thirty is therefore taken to be the minimum number of degrees of freedom for which we have an acceptable level of confidence in an estimate.



Key

X-axis degrees of freedom

Y-axis width of 95 % confidence interval for σ/σ

Figure B.1 — Ratio of the 95 % confidence interval width of a standard deviation estimate versus its degrees of freedom

Annex C
(normative)

Notation and tests

C.1 General

Throughout this document the following notation is used:

— S is the number of samples;

— L is the number of laboratories;

— i is the subscript denoting laboratory number;

— j is the subscript denoting sample number;

— x is an individual test result;

— a is the sum of duplicate test results;

— e is the difference between duplicate test results;

— v is the degrees of freedom.

C.2 Array of duplicate results

Table C.1 shows the array of duplicate results from each of L, laboratories, on S, samples, and corresponding means, m_j.

Table C.1 — Array of duplicate results

Laboratory	Sample			
	1	2	j	S
1	x_{111}	x_{121}	x_{1j1}	x_{1S1}
	x_{112}	x_{122}	x_{1j2}	x_{1S2}
2	x_{211}	x_{221}	x_{2j1}	x_{2S1}
	x_{212}	x_{222}	x_{2j2}	x_{2S2}
i	x_{i11}	x_{i21}	x_{ij1}	x_{iS1}
	x_{i12}	x_{i22}	x_{ij2}	x_{iS2}
L	x_{L11}	x_{L21}	x_{Lj1}	x_{LS1}
	x_{L12}	x_{L22}	x_{Lj2}	x_{LS2}
Total	g_1	g_2	g_j	g_S
Mean	m_1	m_2	m_j	m_S

NOTE If a transformation $y = F(x)$ of the reported data is necessary (see 5.3), then corresponding symbols y_{ij1} and y_{ij2} are used in place of x_{ij1} and x_{ij2}.

C.3 Array of sums of duplicate results

Table C.2 shows the array of sums of duplicate results, of laboratory totals, h_i, and sample totals, g_j.

Table C.2 — Array of sum of results

Laboratory	Sample				
	1	2	j	S	**Total**
1	a_{11}	a_{12}	a_{1j}	a_{1S}	h_1
2	a_{21}	a_{22}	a_{2j}	a_{2S}	h_2
i	a_{i1}	a_{i2}	a_{ij}	a_{iS}	h_i
L	a_{L1}	a_{L2}	a_{Lj}	a_{LS}	h_L
Total	g_1	g_2	g_j	g_S	TOT

In Table C.2 the following applies:

$a_{ij} = x_{ij1} + x_{ij2}$ (or $a_{ij} = y_{ij1} + y_{ij2}$, if a transformation has been used) (C.1)

Similarly the same table can be built for the array of the differences of the results:

$e_{ij} = x_{ij1} - x_{ij2}$ (or $e_{ij} = y_{ij1} - y_{ij2}$, if a transformation has been used) (C.2)

$$g_j = \sum_{i=1}^{L} a_{ij} \tag{C.3}$$

$$h_i = \sum_{j=1}^{S} a_{ij} \tag{C.4}$$

$$m_j = g_j / 2L \tag{C.5}$$

$$TOT = \sum_{i=1}^{L} h_i = \sum_{j=1}^{S} g_j \tag{C.6}$$

If any results are missing from the complete array, then the divisor in the expression for m_j is correspondingly reduced.

C.4 Sums of squares and variances

NOTE See 5.3 for further explanation of the procedure and the variables.

Repeats variance for sample j:

$$d_j^2 = \sum_{i=1}^{L} e_{ij}^2 / 2L \tag{C.7}$$

where L is the number of laboratories.

For sample j where each lab provides two repeats, the degree of freedom for repeats reduces to the number of laboratories, L.

If either or both of a laboratory/sample pair of results is missing, the corresponding term in the numerator is omitted and the factor L is reduced by 1.

Between cells variance for sample *j*:

$$C_j^2 = \left(\sum_{i=1}^{L} \frac{a_{ij}^2}{n_{ij}} - \frac{g_j^2}{S_j} \right) \Big/ (L-1) \tag{C.8}$$

Laboratories variance for sample *j*:

$$D_j^2 = \frac{1}{K_j} \left[C_j^2 + (K_j - 1) d_j^2 \right] \tag{C.9}$$

where

$$K_j = \left(S_j^2 - \sum_{i=1}^{L} n_{ij}^2 \right) \Big/ \left[S_j (L-1) \right] \tag{C.10}$$

n_{ij} is the number of results obtained by laboratory *i* from sample *j*;

S_j is the total number of results obtained from sample *j*;

L is the total number of participating laboratories for sample *j*, equals the number of cells in sample *j* containing at least one result.

Laboratories degrees of freedom for sample *j* is given approximately[11] by

$$v_j = \frac{\left(K_j D_j^2 \right)^2}{\frac{\left(C_j^2 \right)^2}{(L-1)} + \frac{\left[(K_j - 1) d_j^2 \right]^2}{L}} \quad \text{(rounded to the nearest integer)} \tag{C.11}$$

If either or both of a laboratory/sample pair of results is missing, the factor L is reduced by 1.

If both of a laboratory/sample pair of results is missing, the factor $(L - 1)$ is reduced by 1.

C.5 Cochran's test

The largest sum of squares, SS_k, out of a set of n mutually independent sums of squares each based on v degrees of freedom, may be tested for conformity in accordance with

$$\text{Cochran's criterion} = SS_k \Big/ \sum_{i=1}^{n} SS_i \tag{C.12}$$

The test ratio is identical if sum of squares values are replaced by mean squares (variance estimates). If the calculated ratio exceeds the critical value given in Table D.14, then the sum of squares in question, SS_k, is significantly greater than the others with a probability of 99 %. Examples of SS_i include e_{ij}^2 and d_j^2 (see Formula (C.7).

C.6 Hawkins' test

An extreme value in a data set can be tested as an outlier by comparing its deviation from the mean of the data set to the square root of the sum of squares of all such deviations. This is carried out in the form of a ratio. Extra information on variability can be provided by including independent sums of squares into the calculations. These are based on ν degrees of freedom and have the same population variance as the data set in question.

Table C.3 shows the values which are required to apply Hawkins' test to individual samples.

Table C.3 — Values which are required to apply Hawkins' test

Characteristic	Sample			
	1	2	j	S
Number of cells	n_1	n_2	n_j	n_S
Mean of cell means	m'_1	m'_2	m'_j	m'_S
Sum of squares	SS_1	SS_2	SS_j	SS_S

where

n_j is the number of cells in sample j which contains at least one result;

m'_j is the mean of cell means in sample j;

SS_j is the sum of squares of deviations of cell means, a_{ij}/n_{ij}, from the mean of cell means, m'_j, and is given by:

$$SS_j = \sum_{i=1}^{L}\left(\frac{a_{ij}}{n_{ij}} - m'_j\right)^2 \text{ where } m'_j = \frac{1}{n_j}\sum_{i=1}^{L}\left(\frac{a_{ij}}{n_{ij}}\right) \tag{C.13}$$

The test procedure is as follows.

a) Identify the sample, k, and cell mean, a_{ik}/n_{ik}, which has the most extreme absolute deviation $|m'_k - a_{ik}/n_{ik}|$. The cell identified is the candidate for the outlier test, be it high or low.

b) Calculate the total sum of squares of deviations:

$$SS = \sum_{j=1}^{S} SS_j \tag{C.14}$$

c) Calculate the Hawkins' test ratio:

$$B^* = \frac{\left|m'_k - a_{ik}/n_{ik}\right|}{\sqrt{SS}} \tag{C.15}$$

d) Compare the test ratio with the critical value from Table D.15, for $n = n_k$ and extra degrees of freedom ν where

$$\nu = \sum_{j=1}^{S}\left(n_j - 1\right), j \neq k \tag{C.16}$$

e) If B^* exceeds the critical value, reject results from the cell in question (sample k, laboratory i), modify the n_k, m'_k and SS_k values accordingly, and repeat from list item a).

NOTE Hawkins' test applies theoretically to the detection of only a single outlier laboratory in a sample. The technique of repeated tests for a single outlier, in the order of maximum deviation from sample mean, implies that the critical values in Table D.15 do not refer exactly to the 1 % significance level. It has been shown by Hawkins, however, that if $n > 5$ and the total degrees of freedom ($n + \nu$) are greater than 20, then this effect is negligible, as are the effects of masking (one outlier hiding another) and swamping (the rejection of one outlier leading to the rejection of others).

When the test is applied to laboratories averaged over all samples, Table C.3 reduces to a single column. In that case:

n is the number of laboratories = L;

m is the overall mean equal to T/N, where N is the total number of results in the array;

SS is the sum of squares of deviations of laboratory means from the overall mean, and is given by:

$$SS = \sum_{i=1}^{L} \left(\frac{h_i}{n_i} - m \right)^2 \tag{C.17}$$

where n_i is the number of results in laboratory i.

In the test procedure, therefore, identify the laboratory mean, h_i/n_i, that differs most from the overall mean, m.

The corresponding test ratio then becomes:

$$B^* = \frac{|m - h_i/n_i|}{\sqrt{SS}} \tag{C.18}$$

This shall be compared with the critical value from Table D.15 as before, but now with extra degrees of freedom $\nu = 0$. If a laboratory is rejected, adjust the values of n, m and SS accordingly and repeat the calculations.

C.7 Variance ratio test (F-test)

A variance estimate, V_1, based on ν_1 degrees of freedom, can be compared with a second estimate, V_2, based on ν_2 degrees of freedom, by calculating the ratio:

$$F = \frac{V_1}{V_2} \tag{C.19}$$

If the ratio exceeds the appropriate critical value given in Tables D.17 to D.20, where ν_1 corresponds to the numerator (the larger variance estimate) and ν_2 corresponds to the denominator, then V_1 is greater than V_2 at the chosen level of significance.

Annex D
(normative)

Illustration of procedures using ILS results for Bromine Number and statistical tables

The purpose of this annex is to illustrate procedures described from 5.3 to 6.3, using an ILS data set for Bromine Number as reported. This data set has not been pre-screened per 5.2. See Annex F for illustration of procedure in 5.2 using another data set.

D.1 ILS Data for Bromine Number and treatment

ILS data as reported for Bromine Number from 9 laboratories on 8 samples is listed in Table D.1.

Table D.1 — ILS Bromine Number for low boiling samples

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	1,9	64,5	0,80	3,7	11,0	46,1	114,8	1,2
	2,1	65,5	0,78	3,8	11,1	46,5	114,2	1,2
B	1,7	65,4	0,69	3,7	11,1	50,3	114,5	1,2
	1,8	66,0	0,72	3,7	11,0	49,9	114,3	1,2
C	1,8	63,5	0,76	3,5	10,4	48,5	112,4	1,3
	1,8	63,8	0,76	3,5	10,5	48,2	112,7	1,3
D	4,1	63,6	0,80	4,0	10,8	49,6	108,8	1,0
	4,0	63,9	0,80	3,9	10,8	49,9	108,2	1,1
E	2,1	63,9	0,83	3,7	10,9	47,4	115,6	1,3
	1,8	63,7	0,83	3,7	11,1	47,6	115,1	1,4
F	1,8	70,7	0,72	3,4	11,5	49,1	121,0	1,4
	1,7	69,7	0,64	3,6	11,2	47,9	117,9	1,4
G	1,9	63,8	0,77	3,5	10,6	46,1	114,1	1,1
	2,2	63,6	0,59	3,5	10,6	45,5	112,8	0,93
H	2,0	66,5	0,78	3,2	10,7	49,6	114,8	1,1
	1,8	65,5	0,71	3,5	10,7	48,5	114,5	1,0
J	2,1	68,2	0,81	4,0	11,1	49,1	115,7	1,4
	2,1	65,3	0,81	3,7	11,1	47,9	113,9	1,4

Table D.2 lists the values of *m*, *D*, and *d* calculated from the ILS results listed in Table D.1, correct to three significant digits. Corresponding degrees of freedom are in parentheses.

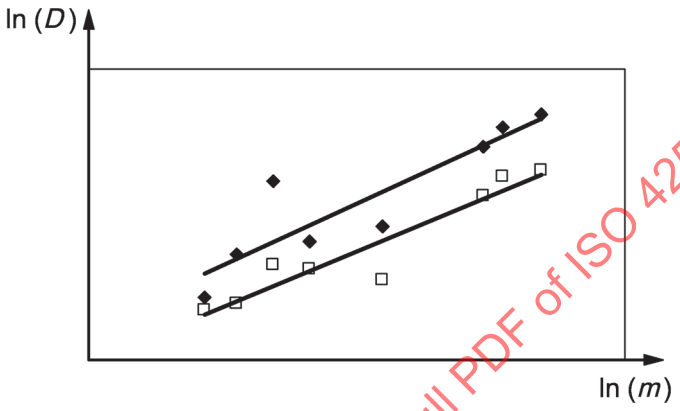
Table D.2 — *m, D, d* statistics of ILS data as reported

Sample number	3	8	1	4	5	6	2	7
m	0,756	1,22	2,15	3,64	10,9	48,2	65,4	114
D	0,066 9 (14)	0,159 (9)	0,729 (8)	0,211 (11)	0,291 (9)	1,50 (9)	2,22 (9)	2,93 (9)
d	0,050 0 (9)	0,057 2 (9)	0,127 (9)	0,116 (9)	0,094 3 (9)	0,527 (9)	0,818 (9)	0,935 (9)

D.2 Transformation determination

This is the procedure as in 5.3.1. Inspection of the numbers in Table D.2 shows that both *D* and *d* increase with *m*, the rate of increase diminishing as *m* increases. A plot of these numbers on log-log paper (i.e. a graph of log *D* and log *d* against log *m*) shows that the points may reasonably be considered as lying about two straight lines (see Figure D.1). As indicated by these plots, based on the form of dependence (2 in Table E.1), the power transformation is most appropriate. The form of the line to be fitted (see Table E.1) using procedure of Clause E.2 is

$$\ln(D) = b_0 + b_1\ln(m) + b_2T + b_3T\ln(m) \quad \text{(D.1)}$$



Key

ln(*m*) logarithm of the sample mean
ln(*D*) logarithm of the standard deviation
◆ laboratories standard deviation
□ repeats standard deviation

Figure D.1 — Log-log plot explaining the power transformation

For a detailed work up of the weighted linear regression, refer to F.4.

The solution from least squared weighted regression is reproduced from Annex F and shown below as Table D.3.

Comparing the t-ratios with the critical 5 % values for 12 degrees of freedom (namely 2,179) given in Table D.16, it can be seen that the slope is significantly non-zero (b1 = 0,638), confirming that a transformation was required. Furthermore, since coefficient b3 does not significantly differ from zero, the slope (and resulting transformation) are the same for both laboratories and repeats standard deviations.

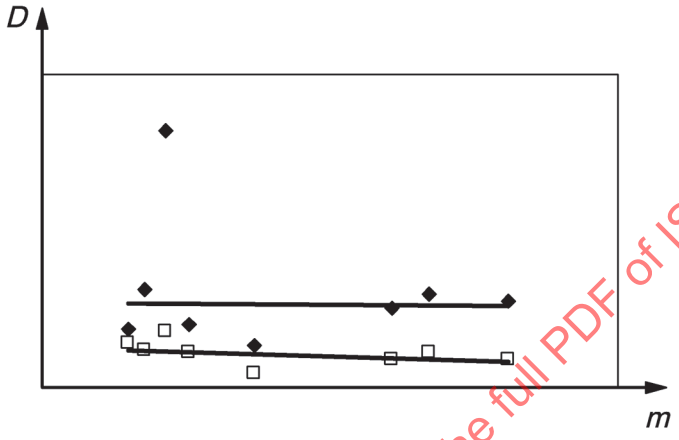
Table D.3 — Solution from least square weighted regression

Fitted variable	Coefficient estimate b_i	Standard error of estimate	*t*-Ratio
Intercept	b_0 = –2,406 4	—	—
ln(*m*)	b_1 = 0,637 73	0,073 59	8,67
Dummy	b_2 = 0,254 96	0,130 52	1,95
Dummy × ln (*m*)	b_3 = 0,028 08	0,047 31	0,59

As the slope b_1 = 0,638 has a standard error of 0,074, the approximate 66 % confidence region of 0,638 ± 0,074 contains the value 2/3. Rounding to this value is, therefore, reasonable, and leads to the convenient transformation

$$y = x^{1/3} \tag{D.2}$$

Having applied this transformation and recalculated sample means and standard deviations, corresponding scatter diagrams (Figure F.2 from Annex F) are reproduced below as Figure D.2. These show uniform levels for both laboratories and repeats standard deviations for all samples except sample 1. Because the data set used in this example has not been pre-screened as per 5.2, the extreme point in Figure D.2 is due to one outlier.



Key

m sample mean (transformed)

D standard deviation (transformed)

◆ laboratories standard deviation (transformed)

□ repeats standard deviation (transformed)

Figure D.2 — Scatter diagram of D and m for transformed values

Hence, the same transformation is appropriate both for repeatability and reproducibility, and is given by Formula (D.3):

$$\int x^{-2/3}\mathrm{d}x = 3x^{1/3} \tag{D.3}$$

Since the constant multiplier may be ignored, the transformation thus reduces to that of taking the cube roots of the reported results (Bromine Numbers). This yields the transformed data shown in Table D.4, in which the cube roots are quoted correct to three decimal places.

Table D.4 — Transformed (Cube root) Bromine Number for low boiling samples

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	1,239	4,010	0,928	1,547	2,224	3,586	4,860	1,063
	1,281	4,031	0,921	1,560	2,231	3,596	4,852	1,063
B	1,193	4,029	0,884	1,547	2,231	3,691	4,856	1,063
	1,216	4,041	0,896	1,547	2,224	3,682	4,853	1,063
C	1,216	3,990	0,913	1,518	2,183	3,647	4,826	1,091
	1,216	3,996	0,913	1,518	2,190	3,639	4,830	1,091
D	1,601	3,992	0,928	1,587	2,210	3,674	4,774	1,000
	1,587	3,998	0,928	1,574	2,210	3,682	4,765	1,032
E	1,281	3,998	0,940	1,547	2,217	3,619	4,871	1,091
	1,216	3,994	0,940	1,547	2,231	3,624	4,864	1,119
F	1,216	4,135	0,896	1,504	2,257	3,662	4,946	1,119
	1,193	4,115	0,862	1,533	2,237	3,632	4,903	1,119
G	1,239	3,996	0,917	1,518	2,197	3,586	4,850	1,032
	1,301	3,992	0,839	1,518	2,197	3,570	4,832	0,976
H	1,260	4,051	0,921	1,474	2,204	3,674	4,860	1,032
	1,216	4,031	0,892	1,518	2,204	3,647	4,856	1,000
J	1,281	4,086	0,932	1,587	2,231	3,662	4,873	1,119
	1,281	4,027	0,932	1,547	2,231	3,632	4,847	1,119

D.3 Cochran Test

This test is to identify a discordant value for absolute difference between a pair of repeat results (procedure in 5.3.3). Continuing with the transformed results in Table D.4 above, the absolute differences (ranges) between transformed repeat results, i.e. of the pairs of numbers in Table D.4 are shown in Table D.5 below.

Table D.5 — absolute differences between transformed repeat results

Laboratory	Sample							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
A	0,042	0,021	0,007	0,013	0,007	0,010	0,008	0,000
B	0,023	0,012	0,012	0,00	0,007	0,009	0,003	0,000
C	0,00	0,006	0,000	0,00	0,007	0,008	0,004	0,000
D	0,014	0,006	0,000	0,013	0,000	0,008	0,009	0,032
E	0,065	0,004	0,000	0,00	0,014	0,005	0,007	0,028
F	0,023	0,020	0,034	0,029	0,020	0,030	0,043	0,000
G	0,062	0,004	0,078	0,00	0,000	0,016	0,018	0,056
H	0,044	0,020	0,029	0,044	0,000	0,027	0,004	0,032
J	0,00	0,059	0,00	0,040	0,000	0,030	0,026	0,000

The largest range is 0,078 for laboratory G on sample 3. The sum of squares of all the ranges is:

$$0,042^2 + 0,021^2 + \ldots + 0,026^2 + 0^2 = 0,043\,9$$

Thus, the ratio to be compared with Cochran's criterion is $\frac{0,078^2}{0,043\,9} = 0,138.$

There are 72 ranges and, as from Table D.14, the criterion for 80 ranges is 0,170 9, this ratio is not significant.

D.4 Hawkins' test

This clause is to identify a discordant pair of results from a laboratory on a particular sample (procedure 5.3.4) The first step is to calculate the deviations of cell means from respective sample means over the whole array. These are shown in Table D.6, in units of the third decimal place. The sum of squares of the deviations is then calculated for each sample. These are also shown in the same table.

Table D.6 — Deviations of cell means from respective sample means over the whole array

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	20	8	14	15	10	48	6	3
B	75	7	20	9	10	47	6	3
C	64	35	3	20	30	4	22	25
D	314	33	18	42	7	39	80	50
E	32	32	30	9	7	18	18	39
F	75	97	31	20	30	8	74	53
G	10	34	32	20	20	61	9	62
H	42	13	4	42	13	21	8	50
J	1	28	22	29	14	8	10	53
Sum of squares	117	15	4	6	3	11	13	17

The cell tested is the one with the most extreme deviation. This was obtained by laboratory D from sample 1. The appropriate Hawkins' test ratio is therefore

$$B^* = \frac{0,314}{\sqrt{0,117 + 0,015 + \ldots + 0,017}} = 0,728\,1$$

The critical value, corresponding to n = 9 cells in sample 1 and ν = 56 [7 samples x (n − 1)] extra degrees of freedom from the other 7 samples, is interpolated from Table D.15 as 0,327 9. The test value is greater than the critical value and so the results from laboratory D on sample 1 are rejected.

As there has been a rejection, the mean, deviations and sum of squares are recalculated for sample 1, and the procedure is repeated. The next cell to be tested is that obtained by laboratory F from sample 2. The Hawkins' test ratio for this cell is:

$$B^* = \frac{0,097}{\sqrt{0,006 + 0,015 + \ldots + 0,017}} = 0,354\,2$$

The critical value corresponding to n = 9 cells in sample 2 and ν = 55 extra degrees of freedom is interpolated from Table D.15 as 0,375 6. As the test ratio is less than the critical value, there are no further rejections.

D.5 Cochran's and variance ratio (*F*) test

This test is executed to reject complete data (from all laboratories) for a sample (procedure in 5). The standard deviations of the transformed results, after the rejection of the pair of results by laboratory D on sample 1, are given in Table D.7 in ascending order of sample mean, correct to three significant digits. Corresponding degrees of freedom are in parentheses.

Inspection shows that there is no visually discordant sample amongst these. It is noted that the standard deviations are now independent of the sample means, which was the purpose of transforming the results.

Table D.7 — Standard deviations of the transformed results of the Bromine Number example

Sample number	3	8	1	4	5	6	2	7
Sample mean	0,910 0	1,066	1,240	1,538	2,217	3,639	4,028	4,851
Laboratories standard deviation	0,027 8 (14)	0,047 3 (9)	0,035 4 (13)	0,029 7 (11)	0,019 7 (9)	0,037 8 (9)	0,045 0 (9)	0,041 6 (9)
Repeats standard deviation	0,021 4 (9)	0,018 2 (9)	0,028 1 (8)	0,016 4 (9)	0,006 3 (9)	0,013 2 (9)	0,016 6 (9)	0,013 0 (9)

The numbers in Table D.8, taken from a different ILS, illustrate the case where an extreme discordant sample is identified by inspection, followed by a formal sample rejection using the variance ratio test in 5.4.

It is clear, by inspection, that the laboratories' standard deviation for sample 93 at 15,26 is far greater than the others. It is noted that the repeats standard deviation in this sample is correspondingly large.

Table D.8 — example from different ILS, illustrating the case of a sample rejection

Sample number	90	89	93	92	91	94	95	96
Sample mean	96,1	99,8	119,3	125,4	126,0	139,1	139,4	159,5
Laboratories standard deviation	5,10 (8)	4,20 (9)	15,26 (8)	4,40 (11)	4,09 (10)	4,87 (8)	4,74 (9)	3,85 (8)
Repeats standard deviation	1,13 (8)	0,99 (8)	2,97 (8)	0,91 (8)	0,73 (8)	1,32 (8)	1,12 (8)	1,36 (8)

Since laboratory degrees of freedom are not the same over all samples, the variance ratio test is used. The variance pooled from all samples excluding sample 93 is the sum of the sums of squares divided by the total degrees of freedom, that is:

$$\frac{\left[\left(8\times 5,10^2\right)+\left(9\times 4,20^2\right)+\ldots+\left(8\times 3,85^2\right)\right]}{(8+9+\ldots+8)}=19,96$$

The variance ratio is then calculated as $(15,26^2)/19,96 = 11,66$.

From Tables D.17 to D.20, the critical value corresponding to a significance level of 0,01/8 = 0,001 25, for 8 degrees and 63 degrees of freedom, is approximately 4. This is less than the test ratio and results from sample 93 shall, therefore, be rejected.

Turning to repeats standard deviations, it is noted that degrees of freedom are identical for each sample and that Cochran's test can therefore be applied. Cochran's criterion is the ratio of the largest sum of squares (sample 93) to the sum of all the sums of squares, that is:

$$2{,}97^2/(1{,}13^2 + 0{,}99^2 + \ldots + 1{,}36^2) = 0{,}510$$

This is greater than the critical value of 0,352 corresponding to $n = 8$ and $v = 8$ (see Table D.14), and confirms that results from sample 93 shall be rejected.

D.6 Estimation of values when one pair of results are missing or rejected

This is the procedure as described in 5.5.2. Since the two results from laboratory D (4th lab) on sample 1 were rejected (see earlier illustration), a_{41} is estimated:

— total of remaining results in laboratory D, $L_1 = 36{,}354$;

— total of remaining results in sample 1, $S_1 = 19{,}845$;

— total of all the results except a_{41}, $T_1 = 348{,}358$.

Also $S' = 8$ and $L = 9$.

Using Formula (D.4),

$$a_{ij} = \frac{1}{(\mathrm{L}-1)(S'-1)} \times (LL_1 + S'S_1 - T_1) \qquad \text{(D.4)}$$

Hence, the estimate of a_{41}, is given by

$$a_{ij} = \frac{1}{(9-1)(8-1)}\left[(9\times 36{,}354) + (8\times 19{,}845) - 348{,}358\right]$$

Therefore $a_{41} = \dfrac{137{,}588}{56} = 2{,}457$

D.7 Hawkins' test for outlying laboratories

The application of this procedure (see 5.6) on the laboratory averages shown in Table D.9 follows exactly that specified in 5.3.4.

The deviations of laboratory averages from the overall mean are given in Table D.9 in units of the fourth decimal place, together with the sum of squares.

Hawkins' test ratio is, therefore,

$$B^* = 263/\sqrt{2219} = 0{,}558$$

Table D.9 — Laboratory averages of the example

Laboratory	A	B	C	D	E	F	G	H	J
Average	2,437	2,438	2,424	2,426[a]	2,444	2,458	2,410	2,427	2,462

[a] Including estimated value.

Table D.10 — Deviations of laboratory averages from the overall mean

Laboratory	A	B	C	D	E	F	G	H	J	SS[a]
Deviation	7	23	125	104	75	220	263	87	254	22,19

[a] Sum of squares.

Comparison with the value tabulated in Table D.15, for n = 9 and ν = 0, shows that this ratio is not significant and, therefore, no complete laboratory rejections are necessary.

D.8 Identification of excessively influential sample(s)- Cook's Distance calculation

As per 5.7.2, the last step prior to proceeding with analysis of variance and calculation of precision estimates is to determine if the selection of transformation function is excessively influenced by one or more samples. Cook's Distance is the recommended statistic for this evaluation. The formula for Cook's Distance is reproduced from 5.7.2 below:

$$\text{Cook's Distance} = \frac{r_i^2}{p} \times \frac{lev_i}{(1-lev_i)} \tag{D.5}$$

where

p = 2 (for regression with slope and intercept);

lev_i is leverage of sample i [see Formula (2)];

r_i is studentized residual of sample i, as follows:

$$r_i = \frac{\varepsilon_i}{s(i)\sqrt{1-lev_i}} \tag{D.6}$$

ε_i in Formula (D.6) is the residual of sample i;

$s(i)$ is the residual mean square obtained from regression with the exclusion of sample i.

$s(i)$ can be calculated by solving Formula (D.7).

$$(n-3)\left[s(i)\right]^2 = (n-2)s^2 - \frac{\varepsilon_i^2}{(1-lev_i)} \tag{D.7}$$

where

$$s^2 = \sum_{i=1}^{n} \varepsilon_i^2 \Big/ (n-2) \tag{D.8}$$

The sample statistics m and D are re-calculated using outlier-free, un-transformed Bromine Number data. A non-weighted linear regression is then performed by regressing ln(D) on ln(m). The outcome from regression are:

$ln(D)\_fit$ = y = 0,702 7 $ln(m)$ – 2,336

where $ln(D)\_fit$ is the regression predicted $ln(D)$.

The outcome from regression and key statistics in Cook's Distance formula [Formula (5)] based on the values as displayed are shown in Table D.11.

Table D.11 — Cook's Distance calculation outcome

	outlier free							
Sample No	*m*	*D*	*ln(m)*	*ln(D)*	*ln(D)_fit*	r_i	lev_i	**Cook's Distance**
3	0,756	0,067	−0,280	−2,705	−2,571	−0,508	0,344	0,07
8	1,22	0,159	0,199	−1,840	−2,231	1,737	0,266	0,55
1	1,91	0,165	0,647	−1,803	−1,912	0,369	0,208	0,02
4	3,64	0,211	1,292	−1,557	−1,454	−0,342	0,151	0,01
5	10,9	0,291	2,389	−1,236	−0,674	−3,053	0,128	0,68
6	48,2	1,496	3,875	0,403	0,383	0,071	0,240	0,00
2	65,4	2,219	4,181	0,797	0,600	0,739	0,283	0,11
7	114	2,934	4,736	1,076	0,995	0,312	0,381	0,03

Inspection of the Cook's Distance column does not reveal any values exceeding 1, thus supporting the conclusion that the transformation was successful.

D.9 Sum of squares

This section describes the forming of the sums of squares for the laboratories x samples interaction sum of squares. Using the transformed results in Table D.4, and, replacing the missing or rejected results by their estimates, the following are the outcome from executing procedures outlined in 6.2.1. These procedures from 6.2.1 are reproduced below:

Mean correction, $M_c = T^2/2L'S'$ (D.9)

where L' is L minus the number of laboratories rejected in 5.6 minus the number of laboratories with no remaining results after rejections in 5.3.4.

$$\text{Samples sum of squares} = \left[\sum_{j=1}^{S'}(g_j^2/2L')\right] - M_c \quad \text{(D.10)}$$

$$\text{Laboratories sum of squares} = \left[\sum_{i=1}^{L'}(h_i^2/2S')\right] - M_c \quad \text{(D.11)}$$

$$\text{Pairs sum of squares} = (1/2)\left[\sum_{i=1}^{L'}\sum_{j=1}^{S'}a_{ij}^2\right] - M_c \quad \text{(D.12)}$$

The laboratories × samples interaction sum of squares, I, is given by:

I = (pairs sum of squares) − (laboratories sum of squares) − (sample sum of squares)

Ignoring any pairs in which there are estimated values,

$$E = \text{repeats sum of squares} = (1/2)\sum_{i=1}^{L'}\sum_{j=1}^{S'}e_{ij}^2 \quad \text{(D.13)}$$

$$\text{Mean correction} = \frac{350{,}815^2}{144} = 854{,}660\,5$$

$$\text{Samples sum of squares} = \frac{22{,}302^2 + 72{,}512^2 + \ldots + 19{,}192^2}{18} - 854{,}660\,5 = 293{,}540\,9$$

Laboratories sum of squares = $\frac{38,992^2+39,020^2+\ldots+39,387^2}{16}-854,660\,5=0,035\ 6$

Pairs sum of squares = (1/2)(2,520^2 + 8,041^2 + ... + 2,238^2) − 854,660 5 = 293,690 8

Repeats sum of squares = (1/2)(0,042^2 + 0,021^2 + ... + 0^2) = 0,021 9

Table D.12 can then be derived.

Table D.12 — Sum of squares of the Bromine Number example

Source of variation	Sum of squares
Samples	293,540 9
Laboratories	0,035 6
Laboratories × samples *(I)*	0,114 3
Pairs	293,690 8
Repeats	0,021 9

If there were no estimated values, the above analysis of variance is exact and 6.2.2 shall be disregarded.

The purpose of performing this approximate analysis of variance is to obtain the minimized laboratories × samples interaction sum of squares, *I*. This is then used as indicated in 6.2.2, to obtain the laboratories sum of squares.

D.10 Forming the sum of squares for the exact analysis of variance

After obtaining the value for the quantity (laboratories × samples interaction sum of squares), denoted by (*I*) above, all the estimated pairs are disregarded and new values of g_j are calculated. The following sums of squares for the exact analysis of variance[10] are formed using formulas from 6.2.2 reproduced below:

$$\text{Uncorrected sample sum of squares} = \sum_{j=1}^{S'} \frac{g_j^2}{S_j} \tag{D.14}$$

where S_j is 2 (L' − number of missing pairs in that sample).

$$\text{Uncorrected pairs sum of squares} = (1/2)\sum_{i=1}^{L'}\sum_{j=1}^{S'} a_{ij}^2 \tag{D.15}$$

The laboratories sum of squares is equal to (pairs sum of squares) − (samples sum of squares) − (the minimized laboratories × samples interaction sum of squares).

$$= (1/2)\left[\sum_{i=1}^{L'}\sum_{j=1}^{S'} a_{ij}^2\right] - \left[\sum_{j=1}^{S'} \frac{g_j^2}{S_j}\right] - I \tag{D.16}$$

Continuing with the Bromine Number example (transformed results), we have:

Uncorrected samples sum of squares = $\frac{19,845^2}{16}+\frac{75,512^2}{18}+\ldots+\frac{19,192^2}{18}=1\,145,183\,4$

Uncorrected pairs sum of squares = $\frac{2,520^2}{2}+\frac{8,041^2}{2}+\ldots+\frac{2,238^2}{2}$ = 1 145,332 9

Therefore, laboratories sum of squares = 1 145,332 9 − 1 145,183 4 − 0,114 3 = 0,035 2.

D.11 Degrees of freedom

This part is following 6.2.3. For the Bromine Number example, there are eight samples and nine laboratories. As no complete laboratories or samples were rejected, then $S' = 8$ and $L' = 9$.

Therefore, laboratories degrees of freedom is $L' - 1 = 8$

Laboratories × samples interaction degrees of freedom: if there had been no estimates, would have been $(9 - 1) \times (8 - 1) = 56$. But since one pair was estimated, the laboratories × samples interaction degrees of freedom = 55.

Repeats degrees of freedom would have been 72 if there had been no estimates. In this case, one pair was estimated, hence repeats degrees of freedom = 71.

D.12 Mean squares and analysis of variance — Test for bias between laboratories

This procedure is described under in 6.2.4. The analysis of variance with degrees of freedoms for the Bromine number example is shown in Table D.13.

Table D.13 — Analysis of variance of the Bromine Number example

Source of variation	Degrees of freedom	Sum of squares	Mean square
Laboratories	8	0,035 2	0,004 400
Laboratories × samples	55	0,114 3	0,002 078
Repeats	71	0,021 9	0,000 308

The ratio M_L/M_{LS} = 0,004 4/0,002 078 has a value 2,117. This is greater than the 5 % critical value obtained from Table D.17, indicating bias between laboratories.

D.13 Expectation of mean squares with estimated values

The formulae and explanations from 6.3.3 are reproduced below:

Laboratories (M_L): $\alpha\sigma_0^2 + 2\sigma_1^2 + \beta\sigma_2^2$

Laboratories × samples (M_{LS}): $\gamma\sigma_0^2 + 2\sigma_1^2$

Repeats (M_r): σ_0^2

where

$\beta = 2\dfrac{(K - S')}{(L' - 1)}$;

K is the number of laboratory × sample cells containing at least one result;

α and γ are computed as follows.

— If there are no cells with only a single estimated result, then $\alpha = \gamma = 1$.

— If there are no empty cells (i.e. every laboratory has tested every sample at least once, and $K = L' \times S'$), then α and γ are both 1 plus the proportion of cells with only a single result.

— If there are both empty cells and cells with only one result, then for each laboratory compute the proportion, p_i, of samples tested for which there is only one result, and the sum, P, of these proportions over all laboratories. For each sample, compute the proportion, q_j, of laboratories that

have tested the sample for which there is only one result, and the sum, Q, of these proportions over all samples. Compute the total number of cells, W, with only one result and the proportion of these among all non-empty cells, W/K. Then:

$$\alpha = 1 + \frac{P - (W/K)}{L' - 1} \tag{D.17}$$

and

$$\gamma = 1 + \frac{W - P - Q + (W/K)}{K - L' - S' + 1} \tag{D.18}$$

Continuing with the Bromine Number example, there are eight samples and nine laboratories, one cell is empty (laboratory D for sample 1), so $K = 71$ and

$$\beta = 2\frac{(71-8)}{(9-1)} = 15{,}75$$

None of the non-empty cells has a single result, so $\alpha = \gamma = 1$.

D.14 Calculation of precision estimates

D.14.1 Repeatability (procedure in 6.3.3.1)

From Table D.13,

Repeatability variance $V_r = 2\sigma_0^2 = 0{,}000\,616$

Repeatability of transformed values $y = t_{71}\sqrt{0{,}000\,616} = 0{,}049\,5$

Repeatability of un-transformed values $x = 3x^{2/3} \times 0{,}049\,5 = 0{,}148x^{2/3}$

D.14.2 Reproducibility (procedure in 6.3.3.2)

The reproducibility variance, V_R, is expressed as

$V_R = 2(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)$

and can be calculated using Formula (D.19):

$$V_R = \frac{2}{\beta}M_L + \left(1 - \frac{2}{\beta}\right)M_{LS} + \left(2 - \gamma + \frac{2}{\beta}(\gamma - \alpha)\right)M_r \tag{D.19}$$

where the symbols are as set out in 6.2.4 and 6.3.3.

The reproducibility estimate is the product of the reproducibility standard deviation and the "t-value", t_v, with appropriate degrees of freedom, v_R, (see Table D.16), corresponding to a two sided probability of 95 %. An approximation[11] to the degrees of freedom of the reproducibility variance, v_R, is given by Formula (D.20).

$$v_R = \frac{V_R^2}{\dfrac{r_1^2}{L' - 1} + \dfrac{r_2^2}{v_{LS}} + \dfrac{r_3^2}{v_r}} \tag{D.20}$$

where

v_{LS} is the degrees of freedom for laboratories × samples;

v_r is the degrees of freedom for repeats;

r_1, r_2 and r_3 are the three successive terms in Formula (D.19), i.e:

$$r_1 = \frac{2}{\beta} M_{\mathrm{L}}$$

$$r_2 = \left(1 - \frac{2}{\beta}\right) M_{\mathrm{LS}}$$

and

$$r_3 = \left(2 - \gamma + \frac{2}{\beta}(\gamma - \alpha)\right) M_{\mathrm{r}}$$

Continuing with the Bromine Number example,

$$\text{Reproducibility variance} = \left(\frac{2}{15{,}75} \times 0{,}004\,40\right) + \left(\frac{13{,}75}{15{,}75} \times 0{,}002\,078\right) + 0{,}000\,308$$

$$= 0{,}000\,559 + 0{,}001\,814 + 0{,}000\,308$$

$$= 0{,}002\,681$$

The degrees of freedom, expressed in the last 4 digits, is:

v_R = 7188/(39 + 60 + 1) = 72 (correct to nearest integer)

Reproducibility of transformed results $y = t_{72}\sqrt{0{,}002\,681} = 0{,}103\,4$

Reproducibility of untransformed result x = 3 $x^{2/3}$ (0,103 4) = 0,310 $x^{2/3}$

Table D.14 — Critical 1 % values of Cochran's criterion for n_r variance estimates and v degrees of freedom

n_r	Degrees of freedom v									
	1	2	3	4	5	10	15	20	30	50
3	0,993 3	0,942 3	0,883 1	0,833 5	0,793 3	0,674 3	0,614 5	0,577 5	0,532 7	0,487 2
4	0,967 6	0,864 3	0,781 4	0,721 2	0,676 1	0,553 6	0,496 4	0,462 0	0,421 3	0,380 8
5	0,927 9	0,788 5	0,695 7	0,632 9	0,587 5	0,469 7	0,416 8	0,385 5	0,348 9	0,313 1
6	0,882 8	0,721 8	0,625 8	0,563 5	0,519 5	0,408 4	0,359 7	0,331 2	0,298 2	0,266 1
7	0,837 6	0,664 4	0,568 5	0,508 0	0,465 9	0,361 6	0,316 7	0,290 7	0,260 6	0,231 6
8	0,794 5	0,615 2	0,520 9	0,462 7	0,422 7	0,324 8	0,283 2	0,259 2	0,231 6	0,205 2
9	0,754 4	0,572 7	0,481 0	0,425 1	0,387 0	0,295 0	0,256 3	0,234 0	0,208 6	0,184 2
10	0,717 5	0,535 8	0,446 9	0,393 4	0,357 2	0,270 4	0,234 2	0,213 5	0,189 8	0,167 3
11	0,683 7	0,503 6	0,417 5	0,366 3	0,331 8	0,249 7	0,215 7	0,196 3	0,174 2	0,153 2
12	0,652 8	0,475 1	0,391 9	0,342 8	0,309 9	0,232 1	0,200 0	0,181 8	0,161 1	0,141 4

NOTE These values are slightly conservative approximations calculated via Bonferroni's inequality[7] as the upper 0,01/n fractile of the beta distribution. If intermediate values are required along the n-axis, they may be obtained by linear interpolation of the reciprocals of the tabulated values. If intermediate values are required along the v-axis, they may be obtained by second order interpolation of the reciprocals of the tabulated values.

Table D.14 *(continued)*

n_r	Degrees of freedom ν									
	1	2	3	4	5	10	15	20	30	50
13	0,624 5	0,449 8	0,369 5	0,322 3	0,290 9	0,216 9	0,186 5	0,169 3	0,149 8	0,131 3
14	0,598 5	0,427 2	0,349 5	0,304 3	0,274 1	0,203 6	0,174 8	0,158 5	0,140 0	0,122 6
15	0,574 7	0,406 9	0,331 8	0,288 2	0,259 3	0,191 9	0,164 5	0,149 0	0,131 5	0,115 0
20	0,479 9	0,329 7	0,265 4	0,228 8	0,204 8	0,149 6	0,127 4	0,115 0	0,101 0	0,087 9
25	0,413 0	0,278 2	0,222 0	0,190 4	0,169 9	0,123 0	0,104 3	0,093 9	0,082 2	0,071 3
30	0,363 2	0,241 2	0,191 4	0,163 5	0,145 5	0,104 6	0,088 5	0,079 4	0,069 4	0,060 0
35	0,324 7	0,213 4	0,168 5	0,143 5	0,127 4	0,091 2	0,076 9	0,069 0	0,060 1	0,051 9
40	0,294 0	0,191 6	0,150 7	0,128 1	0,113 6	0,080 9	0,068 1	0,061 0	0,053 1	0,045 7
45	0,269 0	0,174 0	0,136 4	0,115 8	0,102 5	0,072 7	0,061 1	0,054 7	0,047 5	0,040 9
50	0,248 1	0,159 6	0,124 8	0,105 7	0,093 5	0,066 1	0,055 5	0,049 6	0,043 1	0,037 0
60	0,215 1	0,137 1	0,106 8	0,090 2	0,079 6	0,056 1	0,046 9	0,041 9	0,036 3	0,031 1
70	0,190 3	0,120 4	0,093 5	0,078 8	0,069 5	0,048 7	0,040 7	0,036 3	0,031 4	0,026 9
80	0,170 9	0,107 5	0,083 2	0,070 1	0,061 7	0,043 1	0,036 0	0,032 0	0,027 7	0,023 6
90	0,155 3	0,097 2	0,075 1	0,063 1	0,055 5	0,038 7	0,032 2	0,028 7	0,024 8	0,021 1
100	0,142 4	0,088 8	0,068 5	0,057 5	0,050 5	0,035 1	0,029 2	0,026 0	0,022 4	0,019 1

NOTE These values are slightly conservative approximations calculated via Bonferroni's inequality[7] as the upper $0{,}01/n$ fractile of the beta distribution. If intermediate values are required along the n-axis, they may be obtained by linear interpolation of the reciprocals of the tabulated values. If intermediate values are required along the ν-axis, they may be obtained by second order interpolation of the reciprocals of the tabulated values.

Table D.15 — Critical values of Hawkins' 1 % outlier test for n_R = 3 to 50 and ν = 0 to 200

n_R	Degrees of freedom ν											
	0	5	10	15	20	30	40	50	70	100	150	200
3	0,816 5	0,724 0	0,610 0	0,532 8	0,478 1	0,404 9	0,357 4	0,323 3	0,276 9	0,234 0	0,192 6	0,167 4
4	0,863 9	0,750 5	0,640 5	0,564 4	0,509 4	0,434 5	0,385 0	0,349 2	0,300 0	0,254 1	0,209 6	0,182 4
5	0,881 8	0,757 3	0,653 0	0,579 6	0,525 8	0,451 0	0,401 2	0,364 7	0,314 2	0,266 8	0,220 4	0,192 0
6	0,882 3	0,755 4	0,657 1	0,586 9	0,534 7	0,461 2	0,411 5	0,374 9	0,323 8	0,275 5	0,228 0	0,198 8
7	0,873 3	0,749 3	0,656 7	0,589 8	0,539 4	0,467 6	0,418 4	0,381 9	0,330 7	0,281 9	0,233 7	0,203 9
8	0,859 6	0,740 9	0,653 8	0,590 1	0,541 5	0,471 5	0,423 1	0,386 9	0,335 8	0,286 8	0,238 1	0,207 9
9	0,843 9	0,731 4	0,649 3	0,588 6	0,541 8	0,473 8	0,426 2	0,390 5	0,339 6	0,290 6	0,241 6	0,211 2
10	0,827 4	0,721 3	0,643 9	0,586 1	0,541 1	0,475 0	0,428 3	0,393 0	0,342 6	0,293 6	0,244 5	0,213 9
11	0,810 8	0,711 1	0,638 0	0,582 8	0,539 4	0,475 3	0,429 5	0,394 8	0,344 8	0,296 1	0,246 9	0,216 2
12	0,794 7	0,701 0	0,631 8	0,579 0	0,537 3	0,475 0	0,430 2	0,396 0	0,346 6	0,298 1	0,248 9	0,218 1
13	0,779 1	0,691 0	0,625 4	0,574 9	0,534 7	0,474 2	0,430 4	0,396 8	0,347 9	0,299 7	0,250 7	0,219 8
14	0,764 2	0,681 2	0,618 9	0,570 6	0,531 9	0,473 1	0,430 2	0,397 2	0,348 9	0,301 1	0,252 1	0,221 2
15	0,750 0	0,671 7	0,612 5	0,566 2	0,528 8	0,471 7	0,429 8	0,397 3	0,349 6	0,302 1	0,253 4	0,222 5
16	0,736 4	0,662 5	0,606 1	0,561 7	0,525 6	0,470 1	0,429 1	0,397 2	0,350 1	0,303 0	0,254 4	0,223 6
17	0,723 5	0,653 5	0,599 8	0,557 1	0,522 3	0,468 3	0,428 2	0,396 8	0,350 4	0,303 7	0,255 4	0,224 6
18	0,711 2	0,644 9	0,593 6	0,552 6	0,518 9	0,466 5	0,427 2	0,396 4	0,350 5	0,304 3	0,256 2	0,225 4
19	0,699 6	0,636 5	0,587 6	0,548 0	0,515 5	0,464 5	0,426 0	0,395 8	0,350 6	0,304 7	0,256 9	0,226 2
20	0,688 4	0,628 6	0,581 6	0,543 6	0,512 0	0,462 4	0,424 8	0,395 1	0,350 5	0,305 1	0,257 5	0,226 9
21	0,677 8	0,620 9	0,575 8	0,539 2	0,508 6	0,460 3	0,423 5	0,394 2	0,350 3	0,305 3	0,258 0	0,227 5

Table D.15 *(continued)*

n_R	Degrees of freedom ν											
	0	5	10	15	20	30	40	50	70	100	150	200
22	0,667 7	0,613 4	0,570 2	0,534 8	0,505 2	0,458 1	0,422 1	0,393 4	0,350 0	0,305 5	0,258 4	0,228 0
23	0,658 1	0,606 2	0,564 7	0,530 5	0,501 8	0,455 9	0,420 6	0,392 4	0,349 6	0,305 6	0,258 8	0,228 5
24	0,648 8	0,599 3	0,559 3	0,526 3	0,498 4	0,453 7	0,419 1	0,391 4	0,349 2	0,305 6	0,259 1	0,228 9
25	0,640 0	0,592 5	0,554 0	0,522 1	0,495 1	0,451 5	0,417 6	0,390 4	0,348 8	0,305 6	0,259 4	0,229 3
26	0,631 5	0,586 1	0,549 0	0,518 0	0,491 8	0,449 2	0,416 0	0,389 3	0,348 2	0,305 4	0,259 6	0,229 6
27	0,623 4	0,579 8	0,544 0	0,514 0	0,488 5	0,447 0	0,414 5	0,388 1	0,347 7	0,305 3	0,259 7	0,229 9
28	0,615 6	0,573 7	0,539 2	0,510 1	0,485 3	0,444 7	0,412 9	0,387 0	0,347 1	0,305 1	0,259 9	0,230 2
29	0,608 1	0,567 8	0,534 5	0,506 3	0,482 1	0,442 5	0,411 3	0,385 8	0,346 4	0,304 9	0,260 0	0,230 4
30	0,600 9	0,562 1	0,529 9	0,502 5	0,479 0	0,440 3	0,409 7	0,384 6	0,345 8	0,304 7	0,260 0	0,230 6
35	0,568 6	0,536 1	0,508 6	0,484 8	0,464 1	0,429 4	0,401 6	0,378 5	0,342 1	0,303 1	0,260 0	0,231 2
40	0,541 3	0,513 6	0,489 7	0,468 8	0,450 4	0,419 1	0,393 6	0,372 2	0,338 2	0,301 0	0,259 4	0,231 4
45	0,517 9	0,493 9	0,472 8	0,454 2	0,437 7	0,409 4	0,385 9	0,366 0	0,334 0	0,298 7	0,258 6	0,231 2
50	0,497 5	0,476 4	0,457 7	0,441 0	0,426 0	0,400 2	0,378 5	0,360 0	0,329 9	0,296 2	0,257 5	0,230 8

NOTE The critical values given in Table D.15 are correct to the 4th decimal place in the range n = 3 to 30 and ν = 0, 5, 15 and 30[8]. Other values were derived from the Bonferroni inequality as:

$$B^* = t\sqrt{\frac{(n-1)}{n(n+\nu-2+t^2)}} \tag{D.21}$$

where t is the upper 0,005/n fractile of a t-variate with $n + \nu - 2$ degrees of freedom.

The values so computed are only slightly conservative, and have a maximum error of approximately 0,000 2 above the true value. If critical values are required for intermediate values of n_R and ν_R, they may be estimated by second order interpolation using the square of the reciprocals of the tabulated values. Similarly, second order extrapolation can be used to estimate values beyond n_R = 50 and ν_R = 200.

Table D.16 — Critical values of *t*

Degrees of freedom	Double-sided % significance level						
	50	40	30	20	10	5	1
1	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,700	0,879	1,093	1,372	1,812	2,228	3,165
11	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,691	0,866	1,074	1,341	1,753	2,131	2,947

Table D.16 *(continued)*

Degrees of freedom	Double-sided % significance level						
	50	40	30	20	10	5	1
16	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,683	0,854	1,055	1,310	1,697	2,042	2,750
40	0,681	0,851	1,050	1,303	1,684	2,021	2,704
50	0,680	0,849	1,048	1,299	1,676	2,008	2,678
60	0,679	0,848	1,046	1,296	1,671	2,000	2,660
120	0,677	0,845	1,041	1,289	1,658	1,980	2,617
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,576

D.15 Critical values of *F*

D.15.1 General data representation

See Reference [5] for the source of Tables D.17 to D.20.

Table D.17 — Critical 5 % values of *F*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
3	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,70	8,66	8,62	8,58	8,55	8,54	8,53	8,53
4	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,86	5,80	5,75	5,70	5,66	5,65	5,64	5,63
5	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,62	4,56	4,50	4,44	4,41	4,39	4,37	4,37
6	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	3,94	3,87	3,81	3,75	3,71	3,69	3,68	3,67
7	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,51	3,44	3,38	3,32	3,27	3,25	3,24	3,23
8	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,22	3,15	3,08	3,02	2,97	2,95	2,94	2,93
9	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,01	2,94	2,86	2,80	2,76	2,73	2,72	2,71
10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,85	2,77	2,70	2,64	2,59	2,56	2,55	2,54
15	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,40	2,33	2,25	2,18	2,12	2,10	2,08	2,07
20	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,20	2,12	2,04	1,97	1,91	1,88	1,86	1,84
30	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,01	1,93	1,84	1,76	1,70	1,66	1,64	1,62
50	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,87	1,78	1,69	1,60	1,52	1,48	1,46	1,44
100	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	1,77	1,68	1,57	1,48	1,39	1,34	1,31	1,28

Table D.17 *(continued)*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
200	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	1,72	1,62	1,52	1,41	1,32	1,26	1,22	1,19
500	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,69	1,59	1,48	1,38	1,28	1,21	1,16	1,11
∞	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,67	1,57	1,46	1,35	1,24	1,17	1,11	1,00

Table D.18 — Critical 1 % values of *F*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
3	29,5	28,7	28,2	27,9	27,7	27,5	27,3	27,2	26,9	26,7	26,5	26,4	26,2	26,2	26,1	26,1
4	16,7	16,0	15,5	15,2	15,0	14,8	14,7	14,5	14,2	14,0	13,8	13,7	13,6	13,5	13,5	13,5
5	12,1	11,4	11,0	10,7	10,5	10,3	10,2	10,1	9,72	9,55	9,38	9,24	9,13	9,08	9,04	9,02
6	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,56	7,40	7,23	7,09	6,99	6,93	6,90	6,88
7	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,31	6,16	5,99	5,86	5,75	5,70	5,67	5,65
8	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,52	5,36	5,20	5,07	4,96	4,91	4,88	4,86
9	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	4,96	4,81	4,65	4,52	4,42	4,36	4,33	4,31
10	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,56	4,41	4,25	4,12	4,01	3,96	3,93	3,91
15	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,52	3,37	3,21	3,08	2,98	2,92	2,89	2,87
20	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,09	2,94	2,78	2,64	2,54	2,48	2,44	2,42
30	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,70	2,55	2,39	2,25	2,13	2,07	2,03	2,01
50	4,20	3,72	3,41	3,19	3,02	2,89	2,79	2,70	2,42	2,27	2,10	1,95	1,82	1,76	1,71	1,68
100	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	2,22	2,07	1,89	1,73	1,60	1,52	1,47	1,43
200	3,88	3,41	3,11	2,89	2,73	2,60	2,50	2,41	2,13	1,97	1,79	1,63	1,48	1,39	1,33	1,28
500	3,82	3,36	3,05	2,84	2,68	2,55	2,44	2,36	2,07	1,92	1,74	1,56	1,41	1,31	1,23	1,16
∞	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,04	1,88	1,70	1,52	1,36	1,25	1,15	1,00

Table D.19 — Critical 0,1 % values of *F*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
3	141	137	135	133	132	131	130	129	127	126	125	125	124	124	124	124
4	56,2	53,4	51,7	50,5	49,7	49,0	48,5	48,0	46,8	46,1	45,4	44,9	44,5	44,3	44,1	44,0
5	33,2	31,1	29,8	28,8	28,2	27,6	27,2	26,9	25,9	25,4	24,9	24,4	24,1	23,9	23,8	23,8
6	23,7	21,9	20,8	20,0	19,5	19,0	18,7	18,4	17,6	17,1	16,7	16,3	16,0	15,9	15,8	15,8
7	18,8	17,2	16,2	15,5	15,0	14,6	14,3	14,1	13,3	12,9	12,5	12,2	11,9	11,8	11,7	11,7
8	15,8	14,4	13,5	12,9	12,4	12,0	11,8	11,5	10,8	10,5	10,1	9,80	9,57	9,46	9,39	9,34
9	13,9	12,6	11,7	11,1	10,7	10,4	10,1	9,89	9,24	8,90	8,55	8,26	8,04	7,93	7,86	7,81
10	12,6	11,3	10,5	9,92	9,52	9,20	8,96	8,75	8,13	7,80	7,47	7,19	6,98	6,87	6,81	6,76
15	9,34	8,25	7,57	7,09	6,74	6,47	6,26	6,08	5,53	5,25	4,95	4,70	4,51	4,41	4,35	4,31
20	8,10	7,10	6,46	6,02	5,69	5,44	5,24	5,08	4,56	4,29	4,01	3,77	3,58	3,48	3,42	3,38
30	7,05	6,12	5,53	5,12	4,82	4,58	4,39	4,24	3,75	3,49	3,22	2,98	2,79	2,69	2,63	2,59
50	6,34	5,46	4,90	4,51	4,22	4,00	3,82	3,67	3,20	2,95	2,68	2,44	2,24	2,14	2,07	2,03
100	5,85	5,01	4,48	4,11	3,83	3,61	3,44	3,30	2,84	2,59	2,32	2,07	1,87	1,75	1,68	1,62

Table D.19 *(continued)*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
200	5,64	4,81	4,29	3,92	3,65	3,43	3,26	3,12	2,67	2,42	2,15	1,90	1,68	1,55	1,46	1,39
500	5,51	4,69	4,18	3,82	3,54	3,33	3,16	3,02	2,58	2,33	2,05	1,80	1,57	1,43	1,32	1,23
∞	5,42	4,62	4,10	3,74	3,47	3,27	3,10	2,96	2,51	2,27	1,99	1,73	1,49	1,34	1,21	1,00

Table D.20 — Critical 0,05 % values of *F*

ν_2	ν_1															
	3	4	5	6	7	8	9	10	15	20	30	50	100	200	500	∞
3	225	218	214	211	209	208	207	206	203	201	199	198	197	197	196	196
4	80,1	76,1	73,6	71,9	70,6	69,7	68,9	68,3	66,5	65,5	64,6	63,8	63,2	62,9	62,7	62,6
5	44,4	41,5	39,7	38,5	37,6	36,9	36,4	35,9	34,6	33,9	33,1	32,5	32,1	31,8	31,7	31,6
6	30,4	28,1	26,6	25,6	24,9	24,3	23,9	23,5	22,4	21,9	21,4	20,9	20,5	20,3	20,2	20,1
7	23,5	21,4	20,2	19,3	18,7	18,2	17,8	17,5	16,5	16,0	15,5	15,1	14,7	14,6	14,5	14,4
8	19,4	17,6	16,4	15,7	15,1	14,6	14,3	14,0	13,1	12,7	12,2	11,8	11,6	11,4	11,4	11,3
9	16,8	15,1	14,1	13,3	12,8	12,4	12,1	11,8	11,0	10,6	10,2	9,80	9,53	9,40	9,32	9,26
10	15,0	13,4	12,4	11,8	11,3	10,9	10,6	10,3	9,56	9,16	8,75	8,42	8,16	8,04	7,96	7,90
15	10,8	9,48	8,66	8,10	7,68	7,36	7,11	6,91	6,27	5,93	5,58	5,29	5,06	4,94	4,87	4,83
20	9,20	8,02	7,28	6,76	6,38	6,08	5,85	5,66	5,07	4,75	4,42	4,15	3,93	3,82	3,75	3,70
30	7,90	6,82	6,14	5,66	5,31	5,04	4,82	4,65	4,10	3,80	3,48	3,22	3,00	2,89	2,82	2,78
50	7,01	6,01	5,37	4,93	4,60	4,34	4,14	3,98	3,45	3,16	2,86	2,59	2,37	2,25	2,17	2,13
100	6,43	5,47	4,87	4,44	4,13	3,89	3,70	3,54	3,03	2,75	2,44	2,18	1,95	1,82	1,74	1,67
200	6,16	5,23	4,64	4,23	3,92	3,68	3,49	3,34	2,83	2,56	2,25	1,98	1,74	1,60	1,50	1,42
500	6,01	5,09	4,51	4,10	3,80	3,56	3,36	3,21	2,72	2,45	2,14	1,87	1,61	1,46	1,34	1,24
∞	5,91	5,00	4,42	4,02	3,72	3,48	3,30	3,14	2,65	2,37	2,07	1,79	1,53	1,36	1,22	1,00

D.15.2 Approximate formula for critical values of *F*

Critical values of F for untabulated values of ν_1 and ν_2 may be approximated by second order interpolation from Tables D.17 to D.20.

Critical values of F corresponding to $\nu_1 > 30$ and $\nu_2 > 30$ degrees of freedom and a significance level $100\,(1 - p)$ %, where p is the probability, can also be approximated from the formula:

$$\log_{10}(F) = \frac{A(p)}{\sqrt{b - B(p)}} - C(p)\left(\frac{1}{\nu_1} - \frac{1}{\nu_2}\right) \tag{D.22}$$

where

$$\frac{1}{b} = \frac{1}{2}\left(\frac{1}{\nu_1} + \frac{1}{\nu_2}\right)$$

Values of transformation parameters $A(p)$, $B(p)$ and $C(p)$ are given in Table D.21 for typical values of significance level $100\,(1 - p)$.

Table D.21 — Typical values of formula transformation parameters

100 (1 – p)%	$A(p)$	$B(p)$	$C(p)$
10	1,113 1	0,77	0,527
5	1,428 7	0,95	0,681
2,5	1,702 3	1,14	0,846
1	2,020 6	1,40	1,073
0,5	2,237 3	1,61	1,250
0,1	2,684 1	2,09	1,672
0,05	2,858 0	2,30	1,857

For values of p not given in Table D.21, critical values of F may be obtained by second order interpolation/extrapolation of $\log_{10}(F)$ (either tabulated or estimated from the formula) against $\log_{10}(1 - p)$.

D.16 Critical values of the normal distribution

Critical values, Z, corresponding to a single sided probability, p, or to a double sided significance level $2(1 - p)$, are given in Table D.22 in terms of the "standard normal deviate", where

$$Z = \frac{x - \mu}{\sigma} \tag{D.23}$$

and

where μ and σ are the mean and standard deviation respectively of the normal distribution.

Table D.22 — Critical values of the normal distribution

p	0,70	0,80	0,90	0,95	0,975	0,99	0,995
Z	0,524	0,842	1,282	1,645	1,960	2,326	2,576
$2(1 - p)$	0,60	0,40	0,20	0,10	0,05	0,02	0,01

When p is less than 0,5, the appropriate critical value is the negative of the value corresponding to a probability $(1 - p)$.

Annex E
(normative)

Types of dependence and corresponding transformations

E.1 Types of dependence

The forms of dependence given in Table E.1 are shown graphically in Figures E.1 to E.8. In all cases, K can be any positive constant, and "ln" refers to Napierian logarithms. The form of the line to be fitted includes a dummy variable, T (see F.1), by which it is possible to test for a difference in the transformation as applied to repeatability and reproducibility.

Table E.1 — Forms of dependence

Form of dependence	Transformation	Form of line to be fitted	$\frac{dx}{dy}$	Remarks
1. $D = K(m + B)$ $0 < (m + B)$	$Y = \ln(x + B)$ "log model"	$\ln(D) = b_0 + b_1 \ln(m + B) +$ $\ldots \ldots + b_2T + b_3T \ln(m + B)$ Test: $b_1 = 1$	$(x + B)$	Care shall be taken if $(x + B)$ is small, since rounding becomes critical.
2. $D = Km^B$ $B \neq 1$	$y = x^{1-B}$ "power model"	$\ln(D) = b_0 + b_1 \ln(m) + b_2T$ $+\ldots$ $\ldots + b_3T \ln(m)$ Test: $b_1 \neq 0$	$\frac{X^B}{(1-B)}$	The fitted line will pass through the origin. $B = 1/2$ or 2 are common cases.
3. $D = K(m + B_0)^B$ $B \neq 1$ $B_0 \neq 0$ $0 < (m + B_0)$	$y = (x + B_0)^{1-B}$ "power-with-intercept model"	$\ln(D) = b_0 + b_1 \ln(m + B_0) +$ $\ldots \ldots + b_2T + b_3T \ln(m + B_0)$ Test: $b_1 \neq 0$	$\frac{X^B}{(1-B)}$	The fitted line will not pass through the origin.
4. $d = K\sqrt{\frac{m}{B}\left(1-\frac{m}{B}\right)}$ $0 \leq m \leq B$	$y = \arcsin\sqrt{\frac{x}{B}}$ "arcsin model"	$\ln(D) = b_0 + b_1 \ln[m(B -$ $m)] + \ldots$ $\ldots + b_2T + b_3T \ln[m(B - m)]$ Test: $b_1 = 1/2$	$2\sqrt{x(B-x)}$	This case often arises when results are reported as percentages or qualitatively as "scores". If x is always small, the transformation reduces to $y = \sqrt{x}$, a special case of 2 above.
5. $D = K\frac{m}{B}\left(1-\frac{m}{B}\right)$ $0 \leq m \leq B$	$y = \ln\left(\frac{x}{B-x}\right)$ "logistic model"	$\ln(D) = b_0 + b_1 \ln[m(B -$ $m)] + \ldots$ $\ldots + b_2T + b_3T \ln[m(B - m)]$ Test: $b_1 = 1$	$\frac{x(B-x)}{B}$	This case arises when results are reported on a rating scale of 0 to B. If x is always small, then the transformation reduces to $y = \ln(x)$, a special case of 1 above.
6. $D = K\left(\frac{m^2 + B^2}{B}\right)$ $B > 0$	$y = \arctan\left(\frac{x}{B}\right)$ "arctan model"	$\ln(D) = b_0 + b_1 \ln(m2 + B2)$ $+ \ldots \ldots + b_2T + b_3T \ln(m2 +$ $B+)$ Test: $b_1 = 1$	$\frac{\left(x^2 + B^2\right)}{B}$	The fitted line does not pass through the origin. If B is small, the transformation reduces to $y = 1/x$, a special case of 2 above.

E.2 Transformation procedure

The following steps shall be taken in identifying the correct type of transformation and its parameters B and B_0.

a) Plot laboratories standard deviations, D, and repeats standard deviations, d, against sample means in the form of scatter diagrams. Refer to Figures E.1 to E.8 and identify the type of transformation to be applied (if any).

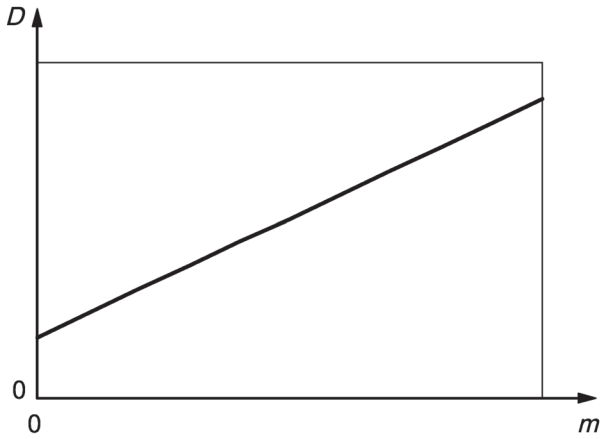
b) With the exception of the power transformations (types 2 and 3 in Table E.1), estimate the transformation parameter, B, from the scatter diagrams. These are known for the arcsin and the logistic transformation (types 4 and 5, respectively, in Table E.1), since B, in both cases, is the upper limit of the rating scale or "score" which defines results. For the log transformation (type 1 in Table E.1), calculate B from the intercept and slope estimated from the scatter diagrams. Similarly, estimate B from the intercept in the case of the arctan transformation (type 6 in Table E.1). In every case, B shall be rounded to give a meaningful value that satisfies the plots for both the laboratories and repeats standard deviations. For the power-with-intercept transformation (type 3 in Table E.1), also estimate the parameter B_0 from the scatter diagrams (but see the Note below).

c) Fit the line specified in Table E.1, corresponding to the transformation in question, according to the computational procedure in Annex F.3 (but see the Note below). For both types of power transformations, coefficient b_1 shall differ significantly from zero and provide an estimate of B, which shall be rounded to a meaningful value. If b_1 is not statistically different from a value of 1, a log transform should be considered. For the arcsin transformation, b_1 shall have a value not significantly different from 0,5. Similarly, b_1 shall not significantly differ from a value of 1 for the logistic, log and arctan transformations.

   In every case, the test specified in Table E.1 shall be applied at the 5 % significance level. Failure of this test implies either that the type of transformation or its parameters B and/or B_0 is/are incorrect. Similarly, the coefficient b_3 shall in every case be tested as zero. Failure in this case implies that the transformation is different for repeatability and reproducibility. In some cases, the presence of outliers can give rise to this difference.

d) If the tests applied in a) to c) above are satisfactory, transform all the results accordingly, recalculate means and standard deviations using transformed results and create new scatter diagrams as in paragraph a) above. These now show a uniform level for laboratories standard deviation, and a uniform (but not necessarily the same) level for repeats standard deviation. Visually confirm that a uniform level for laboratory and repeats standard deviation has been achieved.

NOTE For the power-with-intercept transformation, B and B_0 cannot be estimated together according to the linear least squares technique described in F.3. A non-linear and iterative technique is required instead, such as the simplex procedure of Nelder and Mead[12], necessitating the use of appropriate computer software[13].



Key

m mean

D standard deviation, $D = K(m + B)$, $(m + B) > 0$

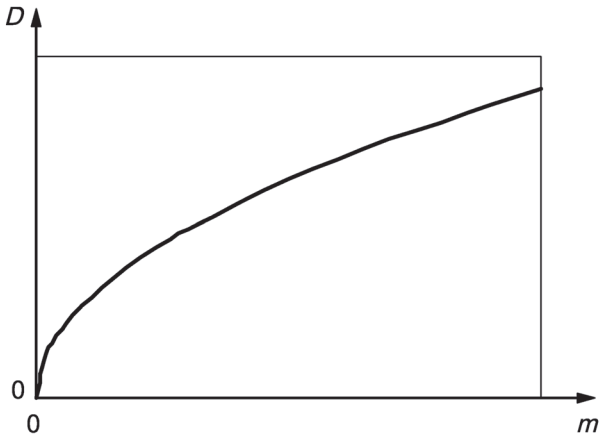
Figure E.1 — Forms of dependence (see Table E.1)

Key

m mean

D standard deviation, $D = Km^B$, $B > 1$

Figure E.2 — Forms of dependence (see Table E.1)

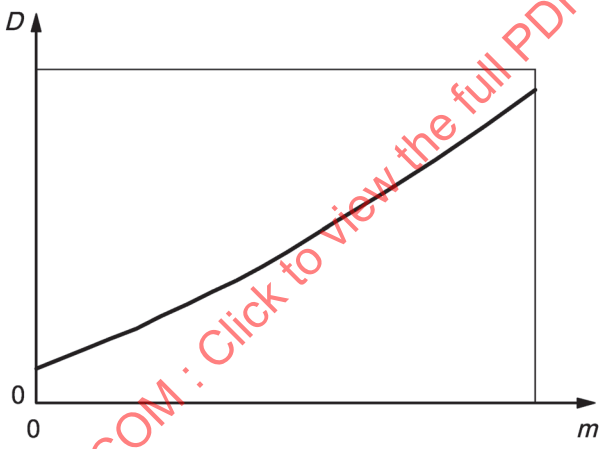


Key

m mean

D standard deviation, $D = Km^B$, $0 < B < 1$

Figure E.3 — Forms of dependence (see Table E.1)

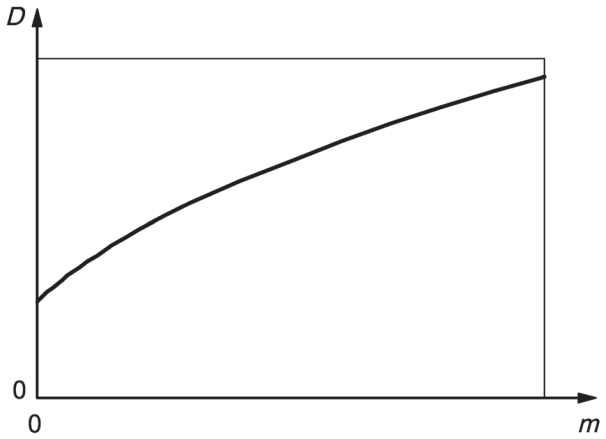


Key

m mean

D standard deviation, $D = K(m + B_0)^B$, $B > 1$, $B_0 \neq 0$

Figure E.4 — Forms of dependence (see Table E.1)

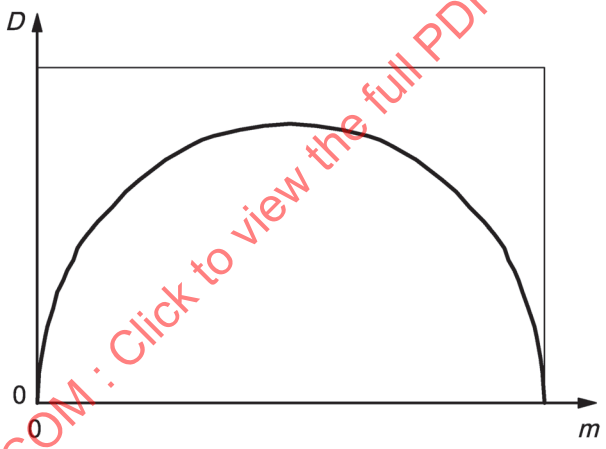


Key

m mean

D standard deviation, $D = K(m + B_0)^B$, $0 < B < 1$, $B_0 \neq 0$

Figure E.5 — Forms of dependence (see Table E.1)

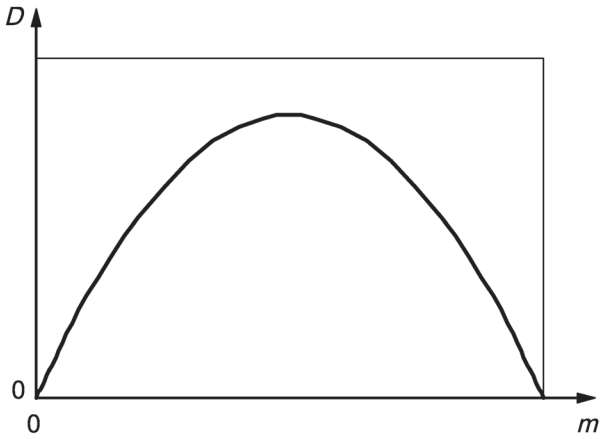


Key

m mean

D standard deviation, $D = K\sqrt{\frac{m}{B}\left(1-\frac{m}{B}\right)}$, $0 \leq m \leq B$

Figure E.6 — Forms of dependence (see Table E.1)

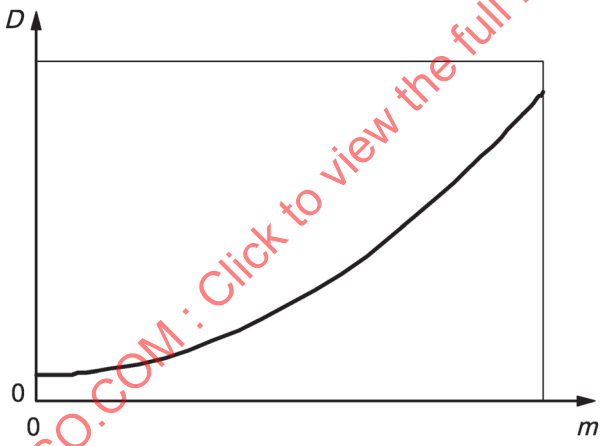


Key

m mean

D standard deviation, $D = K\frac{m}{B}\left(1-\frac{m}{B}\right)$, $0 \le m \le B$

Figure E.7 — Forms of dependence (see Table E.1)



Key

m mean

D standard deviation, $D = K\left(\frac{m^2 + B^2}{B}\right)$, $B > 0$

Figure E.8 — Forms of dependence (see Table E.1)

Annex F
(normative)

Weighted linear regression analysis

F.1 Explanation for the use of a dummy variable

Two different variables, Y_1 and Y_2, when plotted against the same independent variable, X, in general give different linear relationships of the form:

$$Y_1 = b_{10} + b_{11}X$$
$$Y_2 b_{20} + b_{21}X \quad \text{(F.1)}$$

where the coefficients b_{ij} are estimated by regression analysis. In order to compare the two relationships, a dummy variable, T, can be defined such that:

$T = T_1$, a constant value for every observation of Y_1

$T = T_2$, a constant value for every observation of Y_2

$T_1 \neq T_2$

Letting Y represent the combination of Y_1 and Y_2, plot a single relationship,

$$Y = b_0 + b_1X + b_2T + b_3TX \quad \text{(F.2)}$$

where, as before, the coefficients b_i are estimated by regression analysis. By comparing Formulae (F.1) and (F.2), it is evident that

$$b_{10} = b_0 + b_2T_1$$

$$b_{20} = b_0 + b_2T_2 \quad \text{(F.3)}$$

and that therefore

$$b_{10} - b_{20} = b_2(T_1 - T_2) \quad \text{(F.4)}$$

Similarly,

$$b_{11} - b_{21} = b_3(T_1 - T_2) \quad \text{(F.5)}$$

In order to test for a difference between b_{10} and b_{20}, therefore, it is only necessary to test for a non-zero coefficient b_2. Similarly, to test for a difference between b_{11} and b_{21}, test for a non-zero coefficient b_3.

Any non-zero values can be chosen for T_1 and T_2. However, since reproducibility is the basis of tests for quality control against specifications, weighting shall reflect this in the estimation of precision relationships. An "importance ratio" of 2:1 in the favour of reproducibility shall be applied by setting $T_1 = 1$ and $T_2 = -2$, where T_1 refers to the plot of laboratories standard deviation and T_2 refers to the repeats standard deviation.

F.2 Derivation of weights used in regression analysis

In order to account for the relative precision of fitted variables in a regression analysis, weights shall be used that are inversely proportional to the variance of the fitted variables.

For a variable, D, which is an estimate of population standard deviation, σ, based on $\nu(D)$ degrees of freedom, the variance of D is given by

$$Var(D) = \sigma^2/2\nu(D) \tag{F.6}$$

Replacing σ^2 by its estimate D^2, the weight for this variable is approximated by

$$w(D) = 2\nu(D)/D^2 \tag{F.7}$$

It is clear that as standard deviation, D, increases, so does the weight decrease. For this reason, the fitted variable in the weighted regression shall instead be a function of standard deviation which yields weights independent of the fitted variable.

In cases where a function $g(D)$, rather than D itself, is fitted, the variance formula becomes

$$Var[g(D)] = \left(\frac{\delta g}{\delta D}\right)^2 Var(D) \tag{F.8}$$

Hence, for the Napierian logarithm function:

$$Var[\ln(D)] = \frac{1}{D^2} Var(D) = \frac{1}{D^2}\frac{\sigma^2}{2\nu(D)} \tag{F.9}$$

Once again, replacing σ^2 by its estimate D^2, the weight for $\ln(D)$ will be approximated by:

$$w[\ln(D)] = 2\nu(D) \tag{F.10}$$

In relation to laboratories standard deviation D and repeats standard deviation d, therefore, it is necessary to perform regression analysis in terms of $\ln(D)$ and $\ln(d)$, since weighting then takes account only of the amount of data on which the standard deviation was based. A relationship estimated in this way is less dependent on samples that have a high proportion of missing results.

Denoting degrees of freedom as $\nu(D)$ for laboratory standard deviations, D, and $\nu(d)$ for repeats standard deviations, d, formulae for calculating weights then become

$$w[\ln(D)] = 2\nu(D) \tag{F.11}$$

$$w[\ln(d)] = 2\nu(d) \tag{F.12}$$

NOTE Unweighted regression corresponds to weighted regression in which all the weights have a constant value of 1.

F.3 Computational procedure in regression analysis

The following technique gives the best fitting straight line of the form of Formula (F.2) (but see the Note to E.2).

First draw up a table (see Table F.1) giving values of the variables to be plotted in the regression, together with corresponding weights. Functions g_1 and g_2 are always Napierian logarithms corresponding to the transformation in question, as specified in E.2.

Table F.1 — Values of variables to be plotted in the regression and corresponding weights

Sample	Standard deviation function g_1	Sample mean function g_2	Dummy T	Tg_2	Weight
1	$g_1(D_1)$	$g_2(m_1)$	1	$g_2(m_1)$	$2v(D_1)$
2	$g_1(D_2)$	$g_2(m_2)$	1	$g_2(m_2)$	$2v(D_2)$
3	$g_1(D_3)$	$g_2(m_3)$	1	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
S	$g_1(D_S)$	$g_2(m_S)$	1	$g_2(m_S)$	$2v(D_S)$
1	$g_1(d_1)$	$g_2(m_1)$	−2	$-2g_2(m_1)$	$2v(d_1)$
2	$g_1(d_2)$	$g_2(m_2)$	−2	.	$2v(d_2)$
3	$g_1(d_3)$	$g_2(m_3)$	−2	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
S	$g_1(d_S)$	$g_2(m_S)$	−2	$-2g_2(m_S)$	$2v(d_S)$
Symbol	y_i	x_{1i}	x_{2i}	x_{3i}	w_i

Using the symbols defined in Table F.1, the line to be fitted (Formula F.2) becomes:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \tag{F.13}$$

The intercept, b_0, can be eliminated by rewriting this as

$$(y-\bar{y}) = b_1(x_1-\bar{x}_1) + b_2(x_2-\bar{x}_2) + b_3(x_3-\bar{x}_3) \tag{F.14}$$

where

$\bar{y}$, $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$ are weighted means, for example

$$\bar{x}_2 = \frac{\sum_{i=1}^{n} w_i x_{2i}}{\sum_{i=1}^{n} w_i} \tag{F.15}$$

where n_p is the number of points (twice the number of samples) to be plotted.

The least squares solution of Formula (F.14) requires the solution of the set of simultaneous formulae of the form:

$$\begin{aligned} a_{y1} &= a_{11}b_1 + a_{12}b_2 + a_{13}b_3 \\ a_{y2} &= a_{21}b_1 + a_{22}b_2 + a_{23}b_3 \\ a_{y3} &= a_{31}b_1 + a_{32}b_2 + a_{33}b_3 \end{aligned} \tag{F.16}$$

Examples of a_{ij} and a_{yi} elements, in terms of weighted means $\bar{x}_i$, are as follows:

$$a_{22} = \sum w_i (x_{2i} - \bar{x}_2)^2 \qquad a_{23} = \sum w_i (x_{2i} - \bar{x}_2)(x_{3i} - \bar{x}_3)$$

$$a_{y2} = \sum w_i (y_i - \bar{y})(x_{2i} - \bar{x}_2) \qquad a_{yy} = \sum w_i (y_i - \bar{y})^2$$

Having solved the formulae for b_1, b_2 and b_3, calculate the intercept from the weighted means of the variables as

$$b_0 = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2 - b_3\bar{x}_3 \tag{F.17}$$

The coefficient estimates, b_i, can be summarized in tabular form, together with test statistics, as given in Table F.2.

Table F.2 — Coefficient estimates

Fitted variable	Coefficient estimate	Standard error of estimate	*t*-Ratio
Intercept	B_0	e_0	t_0
Sample mean	B_1	e_1	t_1
Dummy	B_2	e_2	t_2
Dummy × mean	B_3	e_3	t_3

In order to complete the table, it is necessary to calculate the standard deviation of the observed y values about the estimated line. This is called the residual standard deviation, *rsd*, and is given by

$$rsd = \sqrt{\frac{1}{n-4}\left(a_w - b_1 a_{y1} - b_2 a_{y2} - b_3 a_{y3}\right)}$$

$$rsd = \sqrt{\frac{1}{n-4}(a_{yy} - b_1 a_{y1} - b_2 a_{y2} - b_3 a_{y3})} \tag{F.18}$$

Standard errors of the estimates then become

$$e_i = rsd\sqrt{c_{ii}} \quad e_i = rsd\sqrt{c_{ii}} \text{ for } i = 1 \text{ to } 3$$

and

$$e_0 = rsd\sqrt{\frac{1}{n} + c_{11}\bar{x}_1^2 + c_{22}\bar{x}_2^2 + c_{33}\bar{x}_3^2 + 2c_{12}\bar{x}_1\bar{x}_2 + 2c_{13}\bar{x}_1\bar{x}_3 + 2c_{23}\bar{x}_2\bar{x}_3} \tag{F.19}$$

where the elements c_{ij} correspond to the inverse of the matrix containing elements a_{ij}.

The *t*-ratios are the ratios $(b_i - K)/e_i$, where K is a constant, and by comparing these to the critical values of t in Table D.16, it is possible to test if coefficient b_i differs from K.If t_i is greater than the critical value corresponding to 5 % significance and $(n - 4)$ degrees of freedom, then the coefficient can be regarded as differing from K. In particular, t_1 identifies an inappropriate slope b_1 and t_3 indicates whether the slope is different for laboratories and repeats standard deviations. Since laboratories standard deviation is generally larger than repeats standard deviation at the same level of sample mean, t_2 in general indicates a non-zero coefficient, b_2.

F.4 Worked example

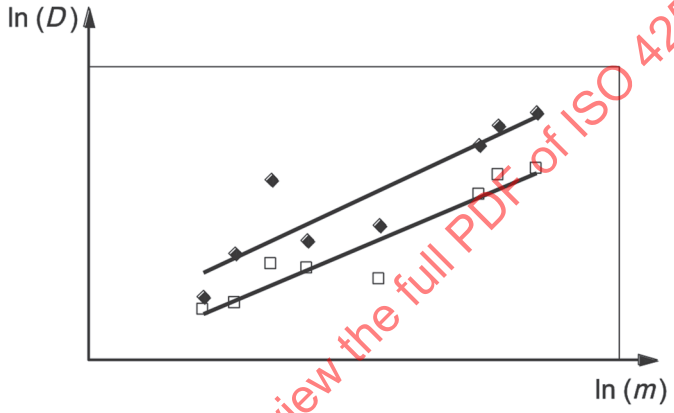
This clause describes the fitting of a power function (form of dependence 2 of Annex E) using weighted linear regression according to the procedure of E.2. Rounded sample means and standard deviations are given in Table 1, based on the bromine data given in D.1.

Scatter diagrams identify the power transformation as appropriate, as indicated by the log-log plot shown in Figure F.1.

It is not necessary to estimate the transformation parameter, B, from Figure F.1, since it is given in the regression analysis that follows.

The form of the line that is being fitted (see Table E.1) is

$$\ln(D) = b_0 + b_1\ln(m) + b_2T + b_3T\ln(m)$$



Key

ln(m) logarithm of the sample mean
ln(D) logarithm of the standard deviation
◆ laboratories standard deviation
□ repeats standard deviation

Figure F.1 — Log-log plot explaining the power transformation

The table of values that are being fitted (see Table F.1) is given in Table F.3.

Table F.3 — Fitted values

Sample	Logarithm of standard deviation	Logarithm of sample mean	Dummy T	Dummy × ln(m)	Weight
1	– 0,315 8	0,765 5	1	0,765 5	16
2	0,796 9	4,180 4	1	4,180 4	18
3	– 2,704 6	– 0,280 2	1	– 0,280 2	28
4	– 1,556 8	1,293 2	1	1,293 2	22
5	– 1,235 8	2,388 8	1	2,388 8	18
6	0,402 9	3,875 5	1	3,875 5	18
7	1,076 2	4,737 8	1	4,737 8	18
8	– 1,840 1	0,197 5	1	0,197 5	18
1	– 2,064 4	0,765 5	–2	– 1,530 9	18