PD ISO/TR 13519:2012

BSI Standards Publication

Guidance on the development and use of ISO statistical publications supported by software

bsi.

...making excellence a habit.™

National foreword

This Published Document is the UK implementation of ISO/TR 13519:2012.

The UK participation in its preparation was entrusted to Technical Committee SS/6, Precision of test methods.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Published by BSI Standards Limited 2013

ISBN 978 0 580 74971 1

ICS 03.120.30

Compliance with a British Standard cannot confer immunity from legal obligations.

This Published Document was published under the authority of the Standards Policy and Strategy Committee on 31 January 2013.

Amendments issued since publication

Amd. No.	Date	Text affected

PD ISO/TR 13519:2012

TECHNICAL REPORT

ISO/TR 13519

First edition
2012-09-01

Guidance on the development and use of ISO statistical publications supported by software

Lignes directrices pour la rédaction et l'application de publications statistiques ISO utilisant des logiciels

Reference number
ISO/TR 13519:2012(E)

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

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.

ISO/TR 13519 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*.

Introduction

This document contains guidance on software to support the development and use of ISO statistical publications.

Aspects covered within this document include

— traceability of data products (figures, tables and other numerical results),

— specification of software,

— categories of support for software, and

— software performance including the use of reference data for testing purposes.

Any references to commercial products of any kind (including but not restricted to software, data or hardware) or links to websites do not imply any approval, endorsement or recommendation by ISO, or any liability.

Guidance on the development and use of ISO statistical publications supported by software

1 Scope

This document provides guidance on the development and use of ISO publications supported by software. The software largely relates to statistical calculations considered by the subcommittees of ISO/TC 69, *Application of statistical methods*, but many other numerical calculations are covered by similar considerations.

In terms of the development of ISO publications, this document gives guidance on the traceability of data products (figures, tables and other numerical results) reproduced in normative-type documents.

In terms of assisting users of ISO publications, this document gives guidance on information that should be included in ISO publications regarding software specification, categories of support for software, and software performance including the use of reference data sets for testing purposes.

Examples are included that illustrate aspects of the guidance provided.

Reference to 'the Committee' in this document relates to the ISO body concerned with developing the relevant ISO publication.

2 Terms and definitions

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

2.1
algorithm
step-by-step procedure describing a specific calculation, given in sufficient detail that it can be implemented in software

2.2
data product
digital data, including graphical figures expressed digitally, in an ISO publication

EXAMPLE 1 A numerical table in an ISO publication.

EXAMPLE 2 A graphical figure expressed digitally as an array of pixels in an ISO publication.

2.3
domain of applicability
set of inputs for which software can be expected to operate in a valid manner

2.4
numerical accuracy
quality of a numerical value expressed in terms of the number of correct decimal digits in absolute or relative form

2.5
problem parametrization
mathematical representation of a problem involving a specific set of defining parameters

NOTE Generally, a problem has more than one parametrization.

EXAMPLE Straight line in two variables: ISO/TS 28037:2010, *Determination and use of straight-line calibration functions*.

A straight line in the variables X and Y can be parametrized as $Y = A_1 + A_2X$ and as $Y = B_1 + B_2(X - 100)$ (and in other ways). The second form is superior to the first when it is to be used as the model in straight-line regression, where X denotes Celsius temperature, taking values close to 100 °C. See ISO/TS 28037:2010[12].

2.6
reference data
data used for testing an item of software for a specific calculation

2.7
reference result
result corresponding to reference data for a specific calculation

NOTE A reference result is provided independently of test software and to a numerical accuracy that is better than that required of the result provided by the test software.

2.8
test result
result produced by test software for reference data for a specific calculation

2.9
test software
software under test

2.10
unit roundoff
distance from 1,0 to the closest floating-point number strictly less than 1,0

NOTE 1 See IEEE 754:2008[1].

NOTE 2 For the purpose of this Technical Report, floating-point numbers constitute part of the system of numerical computation used for providing statistical results.

NOTE 3 All computed values given in this Technical Report were obtained on a computer having unit roundoff of $2{,}22 \times 10^{-16}$.

3 Traceability of data products

3.1 Digital data, including figures, in an ISO document should be traceable to credible sources. Such data are termed “data product” to distinguish them from source or raw data.

3.2 Such traceability (also known as data provenance) requires knowledge of the source data and the processes by which the data product is derived from the source data. These processes are known as data transformations.

NOTE Traceability is vital in scientific work in general, and statistical applications in particular, concerning the credibility of published data.

3.3 In standards work, it is necessary to understand fully the provenance of a derived data product so that those products can be reproduced and updated as necessary.

NOTE ISO 19115:2003[5] relates to traceability considerations. It is applicable particularly to Geographic Information in the form of digital data, but its principles also apply to non-geographic data. ISO 19115:2003 is typically used for large-scale applications requiring databases. For the purposes of this Technical Report, smaller-scale considerations apply, but some concepts are usefully borrowed from this area.

3.4 For any data product, the source data used and the data transformations applied should be recorded in sufficient detail that the data product could be reproduced. Explicit reference should be given to software used for this purpose.

3.5 If the data product is a graphical figure, traceability relates to normal visual resolution. Otherwise, traceability relates to the number of decimal digits quoted in the ISO publication.

EXAMPLE ISO Technical Specification with data products provided by MATLAB and LaTeX.

Each numerical table in ISO/TS 28037:2010 was produced using a MATLAB script that generated LaTeX source, which was combined with other LaTeX source to provide a PDF version of ISO/TS 28037:2010. The table is a data product that is provided by a transformation (implemented in the MATLAB script) of a data set that corresponded to the data points — the source data — or data derived similarly from the source data. Any change at the drafting stage in the source data or transformations was hence reflected in the resulting data product. Consequently no human transcription error was incurred and the numerical tables (and accompanying text) would be reproducible through re-use of this process.

The source data and details of the transformations were recorded in ISO/TS 28037:2010.

NOTE 1 ISO Central Secretariat can provide a suitable website to hold traceability records including software.

NOTE 2 MATLAB is the trade name of a product supplied by The MathWorks Inc.

3.6 The rounding rules given in Annex B of ISO 80000-1:2009[6] apply to the expression of decimal numbers.

4 Specification of software

An ISO publication should contain a specification of software that would materially assist in using that publication. The software of concern largely relates to statistical calculations. The specification would include details of

a) the inputs to the software,

b) the outputs from the software,

c) how the outputs are obtained from the inputs,

d) how the inputs and outputs and relevant intermediate values relate to the various (sub)clauses and quantities in the ISO publication, and

e) the performance of the software, particularly in terms of the numerical accuracy required in the results produced. Also see 6.2.

EXAMPLE 1 ISO 13528:2005, *Statistical methods for use in proficiency testing by interlaboratory comparisons*

Subclause 8.6 of ISO 13528:2005[4] is concerned with producing a graph of repeatability standard deviation for each laboratory involved in a proficiency testing scheme against the corresponding average for the laboratory. It also specifies the construction of a confidence region corresponding to a particular significance level 1 – P under the assumption of normality. The generic specification in this case for such a graph is as follows:

a) Software inputs

p number of participating laboratories,

n number of replicate indication values,

x_i average indication value for laboratory i, $i = 1, \ldots, p$,

s_i within-laboratory standard deviation for laboratory i, $i = 1, \ldots, p$, and

P significance level (for example, 0,05).

b) Software outputs

x x-coordinates used in defining the boundary of the region having a significance level of 100P %, and

s corresponding y-coordinates.

c) Outputs obtained from inputs using

values of x lying in the interval $\bar{X} - (\nu / n)^{1/2}\bar{S}$ to $\bar{X} + (\nu / n)^{1/2}\bar{S}$, where $\nu = \chi^2_{2;1-P}$, and

values of s, corresponding to those of x, equal to $\bar{S}\exp\left(\pm(2n-2)^{-1/2}\left\{ \nu - \left[n^{1/2}(x-\bar{X})/\bar{S} \right]^2 \right\}^{1/2} \right)$, where

$\bar{X} = x^*$, the robust average of $x_1, \ldots, x_p$, and $\bar{S} = s^*$, the robust pooled value of $s_1, \ldots, s_p$ [see d)].

d) Inputs and outputs and relevant intermediate values relate to the various (sub)clauses and quantities as follows:

Items in a) and b) and how items in b) are obtained from those in a) are described in subclause 8.6.1 of ISO 13528:2005,

x^* is calculated by Algorithm A in Annex C.1 of ISO 13528:2005,

s^* is the value of w^* calculated by Algorithm S in Annex C.2 of ISO 13528:2005, and

the expression for s is given in Formula (41) of ISO 13528:2005.

e) Performance of the software

Standard deviations are expressed to a prescribed number of significant decimal digits (usually one or two), and related averages expressed to the same number of decimal places, for inputs lying within a stated domain of applicability.

In any particular instance, numerical values for the inputs (p, n, the x_i, the s_i and P) would be provided. Figure 11 in ISO 13528:2005 is a specific instance of a graph produced by plotting the points (x_i, s_i), $i = 1, \ldots, p$. It also shows the boundaries of confidence regions corresponding to a significance level P, for P = 0,001, 0,01, 0,05.

EXAMPLE 2 ISO/TS 28037:2010, Determination and use of straight-line calibration functions

Clause 9 of ISO/TS 28037:2010[12] is concerned with obtaining estimates a and b of the parameters A and B in a straight-line calibration function $Y = A + BX$ relating variables X and Y. Measured values (data points) x_i and y_i of X and Y are available, as are associated standard uncertainties $u(y_i)$ and covariances cov(y_i, y_j) associated with each pair (y_i, y_j), $i = 1, \ldots, m$, $j = 1, \ldots, m, j \neq i$. Uncertainties associated with the x_i are regarded as negligibly small in this clause of ISO/TS 28037:2010, as are other covariances. The specification in this case is as follows:

a) Software inputs

m	number of data points,
x_i	measured value of X-coordinate of the ith data point, $i = 1, \ldots, m$,
y_i	measured value of Y-coordinate of the ith data point, $i = 1, \ldots, m$,
$u(y_i)$	standard uncertainty associated with y_i, $i = 1, \ldots, m$, and
cov(y_i, y_j)	covariance associated with pair (y_i, y_j), $i = 1, \ldots, m, j = 1, \ldots, m, j \neq i$.

b) Software outputs

a	estimate of A,
b	estimate of B,
$u(a)$	standard uncertainty associated with a,
$u(b)$	standard uncertainty associated with b, and
cov(a, b)	covariance associated with a and b.

c) Outputs obtained from inputs using algorithm given in subclause 9.2.2 of ISO/TS 28037:2010.

NOTE A software implementation is provided at http://standards.iso.org/iso/ts/28037/ .

d) Inputs and outputs and relevant intermediate values relate to the various (sub)clauses and quantities as follows:

Items in a) and b) and how items in b) are obtained from those in a) are described in subclauses 9.1 and 9.2 of ISO/TS 28037:2010.

e) Performance of the software

The software obtains results having adequate numerical accuracy for practical calibration problems, by the use of stable numerical methods (see 6.2). The software reproduces the data products (figures, tables and other numerical results) in ISO/TS 28037:2010 in accordance with Clause 3.

5 Software support

5.1 Support for software implementation

5.1.1 The Committee should consider any software aspects relevant to an ISO publication with which it is concerned and the consequent categories of support to be provided. Categories of support for the task in hand relate to the following:

0 no support;

1 general-purpose software provided by a third party;

2 specific software provided by a third party;

3 stand-alone software accompanying the ISO publication, that is, software that can be run by the user of the ISO publication;

4 sample code included in the ISO publication for extension or generalization by users or third parties;

5 an algorithm (step-by-step procedure) for implementation as software by users or third parties.

The categories of support to be provided in any particular case should be stated in the Scope of the relevant ISO publication, and the Committee should be prepared to provide such support.

NOTE 1 The “third party” is the provider of the software, the first party being the user of the ISO publication, and the second party the Committee involved in preparing the publication.

NOTE 2 Instances of 2 and 3 are given in 5.1.3 Example.

5.1.2 The following information should be given in the relevant ISO publication where appropriate:

a) domain of applicability of the software;

b) numerical accuracy expected in the outputs from application of the software. A measure of the sensitivity of the outputs to changes in the inputs can also be given. Also see 6.2.4;

NOTE The numerical accuracy in general depends on the input data and is specific to that data.

c) any results to be provided by the software in addition to the outputs. Such results may be those of intermediate calculations;

d) execution time and memory requirements of the software.

EXAMPLE The time to compute a median for N input data is comparable to that required to sort N floating-point numbers using a fast sorting algorithm.

The Committee should consider which of a) to d) apply to the relevant ISO publication, and in terms of these the appropriate categories of support in 5.1.1.

NOTE 1 Since execution time in d) depends on many factors, including hardware and programming language, this time can be stated in a relative sense.

NOTE 2 Software performance issues are considered in 6.2.

5.1.3 For each relevant task described in an ISO publication, the Committee should decide and document in the publication whether the task can be implemented by any of the following means:

A Any software capable of carrying out the task required;

B As A, but to a particular level of performance;

C Only software that implements a specified algorithm;

D Only a specific software implementation (for example named third-party software, or an executable software application provided with the ISO publication).

EXAMPLE 1 Illustration of category A: ISO/TR 12845:2010, Selected illustrations of fractional factorial screening experiments.

ISO/TR 12845:2010[9] utilizes statistical software of current vintage for experimental data analysis in six detailed examples. In each example a software product is used to produce output to aid in the understanding of the results for the experimental data considered. Four different software products in all were used:

a) JMP, the trade name of a product supplied by SAS Inc.;

b) Design-Expert, the trade name of a product supplied by Stat-Ease, Inc.;

c) MINITAB, the trade name of a product supplied by Minitab Inc.;

d) R, a free software environment for statistical computing and graphics, which compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.

NOTE 1 This information is given for the convenience of users of ISO/TR 12845:2010 and does not constitute an endorsement by ISO of the product named.

NOTE 2 Other statistical software products exist to analyse experimental data of the type considered.

NOTE 3 Ideally, particularly for critical applications, it may be useful to compare the results produced by more than one software product to help establish whether the results are fit for purpose. The results may differ for a variety of reasons and users may seek to understand the reasons for the differences.

NOTE 4 ISO/TR 14468:2010[10] and ISO/TR 29901:2007[11] also make use of JMP and Minitab for analysing experimental data.

EXAMPLE 2 Illustration of categories A to D: modelling of data by a straight-line function.

Consider the problem of modelling data points x_i, y_i, $i = 1, \ldots, m$, by a straight line function using the method of least squares, where (a) the standard uncertainties $u(y_i)$ associated with the specified values of y_i are available, (b) the uncertainties associated with the specified values of x_i are negligible, and (c) there is no covariance associated with any pair of data values.

For this problem, an example of A is general-purpose software for solving a linear least-squares problem of the form

$$\min \sum_{i=1}^{m} f_i^2(A_1, \ldots, A_n),$$

where the f_i depend linearly on $A_1, \ldots, A_n$, and the minimization is carried out with respect to $A_1, \ldots, A_n$.

An example of B is software that is general purpose as immediately above, but implements a numerically stable algorithm, thus ensuring that floating-point rounding errors do not adversely influence the results of the calculation. In particular, the software is to achieve a prescribed performance in terms of the numerical accuracy delivered (6.2), at least for any examples in the relevant ISO publication. See Example 1 e) in Clause 4.

An example of C is a software implementation of an algorithm specifically designed for solving the problem

$$\min_{A,B} \sum_{i=1}^{m} \left[\frac{y_i - A - Bx_i}{u(y_i)} \right]^2,$$

where A and B are the intercept and gradient of the straight line, respectively, and $y_i - A - Bx_i$ is the residual deviation of the model $A + BX$ at $X = x_i$.

An example of D is the software provided to accompany Clause 6 of ISO/TS 28037:2010[12]. See Clause 4, Example 2.

5.1.4 A particular part or parts of a large software package might be relevant and, if so, would need to be identified.

5.1.5 The results provided by the software can be influenced by

a) the version of the software,

b) the version of the software environment,

c) the version of the programming language,

d) the version of the operating system, and

e) the computer hardware.

As far as possible, a choice of a) to e) should be made that minimizes these influences. These aspects also relate to the maintenance of software considered in 6.3.

5.2 Nature of supporting software

5.2.1 Software that supports statistical calculations can take the following forms:

a) free, open-source software;

b) commercial, open-source software;

c) free, closed-source software;

d) commercial, closed-source software.

NOTE 1 Free open-source software is software that is developed, tested and modified through public collaboration and distributed with the objective that it should be shared with others.

NOTE 2 Commercial open-source software is software that contains elements of free open-source software, but possibly with limited functionality.

NOTE 3 Considerations are that (a) free software products are more likely to be employed by users of ISO publications, and (b) open-source software is visible to inspection, and generally has had more attention paid to its presentation.

5.2.2 If transparency, in addition to traceability, is a primary consideration, the (decreasing) order of preference is a) to d) in 5.2.1.

5.2.3 An overriding consideration is practical implementation of the software in the user community.

5.2.4 Consideration should be given to the particular software brands already in use among the intended audience for the ISO publication.

NOTE There is little value in recommending or requiring software that would not operate on the platforms used by the majority of users of the ISO publication under consideration.

5.3 Software languages

5.3.1 In choosing the language in which to provide the sample code in support category 4 (see 5.1.1), consideration should be given to the following aspects:

a) Portability: compilers for languages such as C, C++, Java and Fortran are widely available for many different software platforms;

b) Stability: languages that are stable over time have advantages. For example, programs that have run satisfactorily in early versions of Fortran will generally still so run in more recent versions of the language;

c) Standards compliance: languages specified by ISO are preferred;

d) End-user application: programming languages that cannot be readily installed by the target audience do not materially assist implementation of the ISO publication; proprietary programming languages may be more easily installed by a particular audience.

5.3.2 If appropriate, sample code in more than one language should be provided.

5.3.3 An overriding consideration is the extent to which the code provided assists implementation of the ISO publication.

6 Quality

6.1 General

6.1.1 Software that supports statistical calculations should be developed according to a Software Quality System. ISO 9126-1:2001[7] provides a model for evaluating software quality. See Figure 1. Its provisions relate to a number of the subclauses in this Technical Report.

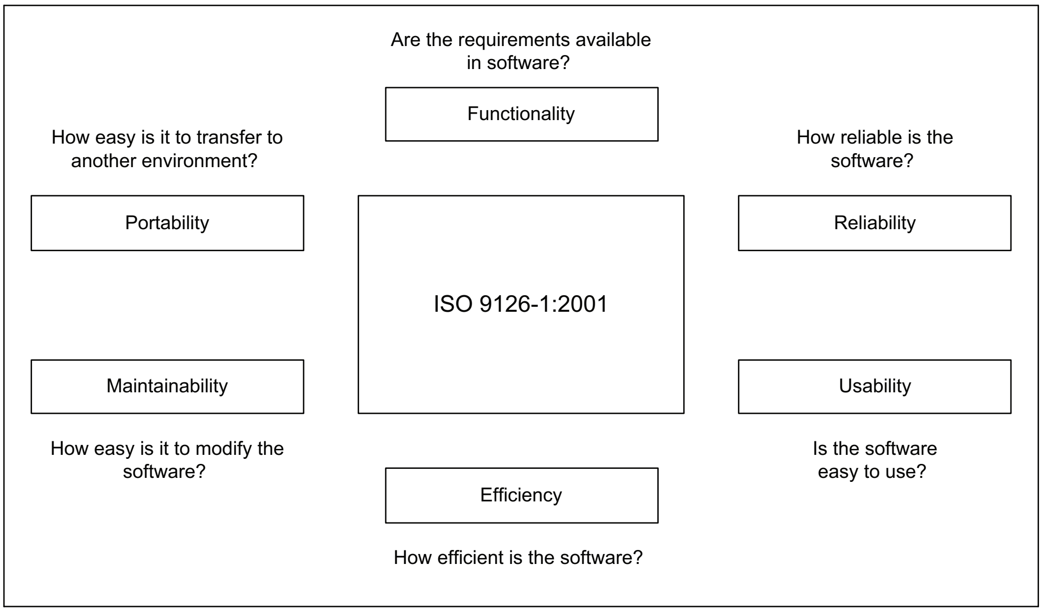

Figure 1 — ISO 9126-1:2001 model for evaluating software quality

The software used should be validated for its quality by using reference data or another comparable facility.

6.1.2 Sample code provided explicitly in the ISO publication should follow code readability guidelines[15]. In particular, for the purpose addressed here, it should

— include adequate comments that relate explicitly to appropriate subclauses of the ISO publication, and

— use meaningful identifier names, preferably relating to the terminology of the ISO publication.

6.1.3 ISO/IEC 18019:2004[8] can be consulted regarding the provision of user documentation.

6.2 Software performance

6.2.1 Examples given in an ISO publication

To obtain the numerical results in an example in an ISO publication, it may in many cases be sufficient to use commercial software (for example, Microsoft Excel, MATLAB, Minitab, SAS). When the calculations only include basic statistics that can be performed by a single command (for example, performing a t-test or an analysis of variance), it should be sufficient to refer to the software used. However, it should not be assumed that all commercial software would provide results that are fit for purpose. The calculation of, for example, the standard deviation of a set of numbers (see page 109 of Reference [13]; also see Example 1 in 6.2.4) might not be carried out sufficiently well by all software available for this task. In such cases, the ISO publication should identify a specific item of software that is fit for purpose. In some cases, suitably pre-processing the data might render it more suitable for applying certain items of software. When appropriate the results should be suitably post-processed.

EXAMPLE Centring and scaling in regression.

Consider transforming the independent variable X in a regression problem by centring and scaling:

$$X' = \frac{X - \bar{x}}{s}.$$

Here $\bar{x}$ is a location value and s is a scale value chosen such that the new variable X' takes on values for the data that are bounded in magnitude by unity. Particularly if the original data values span a wide interval, the use of the transformed values gives rise to a more stable numerical calculation.

For a multiple regression problem, a transformation can be applied to each component in turn. For further information, see Reference [13].

6.2.2 Reference data

6.2.2.1 In order to assess the fitness for purpose of relevant software, an approach based on the use of reference data are recommended. Full details of this approach are given in Reference [13]. Such an approach can be used to obtain an indication of the degree of consistency of the software with the provisions of the ISO publication. In outline, the approach is based on the following points.

a) The software under test, the test software, is regarded as a ‘black box’. Thus, the only interaction with the software is the input data received by the software and the corresponding results returned.

b) Correctly operating test software will produce results to the numerical accuracy required for all practical inputs within the domain of applicability of the software.

c) The testing process is designed to yield information on the extent to which b) is achieved.

d) A suitable suite of reference data and corresponding reference results is available. The suite should cover the practical inputs in b) to the extent possible.

e) These reference results are to be provided to a numerical accuracy at least as great as that to which the data products in the ISO publication are given or are to be given.

f) The test results, namely the results provided by the test software for reference data, are compared with the reference results, to check the suitability of the software.

6.2.2.2 The test software should at least relate to any examples given in the ISO publication, and be capable of reproducing the results therein.

6.2.2.3 The reference data and corresponding reference results indicated in 6.2.2.1 d) should be mutually compatible to the extent required. Suppose the test software is required to produce test results for reference data to a stipulated numerical accuracy. The corresponding reference results should be available to a numerical accuracy that is at least as great as that of the test results. Ideally, the increased numerical accuracy should correspond to one or two decimal (guard) digits.

6.2.2.4 The comparison process indicated in 6.2.2.1 f) involves determining a measure of departure of the test results from the reference results. This departure may be assessed in absolute or relative terms as appropriate.

NOTE NIST provides reference data sets[20] with certified values for a variety of statistical methods.

6.2.3 Comparing results from several software implementations

When reference data sets are unavailable, it may be possible to assess the fitness for purpose of software under consideration by comparing the results it produces with the results from other software implementations, for the same inputs. Such an examination at best can only demonstrate consistency of two or more software implementations. It does not confirm correctness. However, it should be possible to have greater confidence in the software under consideration if it compares favourably with other software that has been independently developed.

NOTE Data for an application currently under consideration can be used in place of reference data. The comparison then relates to the fitness of the software for that particular application and that particular data.

6.2.4 Numerical analysis issues

Numerical analysis issues are relevant in many cases of statistical software. For instance, there may be more than one way of expressing an algorithm or formula, which are mathematically equivalent, but when implemented in floating-point arithmetic have different numerical behaviour.

EXAMPLE 1 Average and standard deviation of repeated indication values in the context of precise measurement.

The arithmetic mean $\bar{x}$ and standard deviation s of m values x_i, $i = 1, \ldots, m$, are given by

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i , \quad s^2 = \frac{1}{m-1}\sum_{i=1}^{m}\left(x_i - \bar{x}\right)^2 . \qquad (1)$$

Apart from exceptional cases, most software for computing $\bar{x}$ from given x_i will produce a result that has relative numerical accuracy approaching the unit roundoff and hence be acceptable in all application areas. The quality of software for computing s varies across implementations.

Consider the following indication values of the mass in kg of a nominally 1 kg weight, obtained with a high-precision balance such as at a national metrology institute:

1,000 000 008 1,000 000 012 1,000 000 009 1,000 000 009 1,000 000 011 (2)

Use of the functions `average` and `stdev` in Microsoft Excel 2000 (version 9.0 6926 SP-3), stating the results to 9 decimal places, gave

$\bar{x}$ = 1,000 000 010 kg, s = 0,000 000 021 kg.

To 9 decimal places, the average is correct, but s differs greatly in a relative sense from its proper value of 0,000 000 002 kg.

The potentially numerically unstable formula[13]

$$s^2 = \frac{1}{m-1}\left(\sum_{i=1}^{m} x_i^2 - m\bar{x}^2\right), \qquad (3)$$

but mathematically equivalent to that in formulæ (1), is used by that version of Microsoft Excel to carry out such computations. The form (3) suffers from subtractive cancellation for data sets with small coefficient of variation $s/\bar{x}$. The standard deviation s given in formulæ (1) is much more numerically stable in such cases.

EXAMPLE 2 As Example 1, but for indication values typical of those provided by an industrial calibration laboratory.

Consider the following indication values of mass in kg typically provided by an industrial calibration laboratory. These values have associated uncertainty three orders of magnitude larger than in Example 1.

$$1{,}000\,008 \quad 1{,}000\,012 \quad 1{,}000\,009 \quad 1{,}000\,009 \quad 1{,}000\,011 \tag{4}$$

Use of the same Microsoft Excel functions, stating the results to 6 decimal places, gives

$$\bar{x} = 1{,}000\,010 \text{ kg}, \qquad s = 0{,}000\,002 \text{ kg},$$

which are correct to the six decimal places given.

NOTE 1 Data (4) were taken as $1 + 10^3(x_i - 1)$, where the x_i denote the data (2) in Example 1.

NOTE 2 The use of unstable Formula (3) is adequate for data of the form (4) and comparable data when using a computer with unit roundoff of $2{,}22 \times 10^{-16}$ or smaller. Nevertheless, it is recommended that formulæ (1) are always used.

EXAMPLE 3 Angle between two vectors.

Consider the calculation of the angle θ between two unit vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.[14] Mathematically equivalent formulæ for evaluating angle $\theta =$ angle AOB (Figure 2) are

$$\theta = \cos^{-1}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}) \tag{5}$$

and

$$\theta = 2\sin^{-1}\left\|\frac{\boldsymbol{b}-\boldsymbol{a}}{2}\right\|. \tag{6}$$

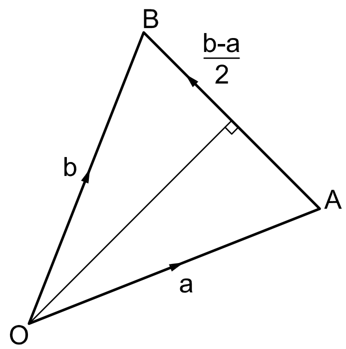

Figure 1 — Angle between two unit vectors

Formula (5) is simpler and seemingly appropriate for the task. The difference in the behaviour of the formulæ especially for an application in dimensional metrology (below), is, however, appreciable.

Consider the vectors to be very nearly parallel, as in ISO 10360-6:2001[3], that is

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 1 - \delta$$

for a positive value δ that is very much smaller than unity. Formula (5) gives, approximately,

$$\cos\theta = 1 - \frac{\theta^2}{2} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 1 - \delta,$$

that is,

$$\theta = \sqrt{2\delta}.$$

In stating the results of floating-point error analyses, the unit roundoff, denoted here by η (definition 2.10), is used. For IEEE arithmetic[1][2], as used on many scientific computers, η takes the value $2{,}22 \times 10^{-16}$. Since the smallest non-zero δ for which $1 - \delta$ is distinguishable from unity on the computer is $\eta/2$, the smallest positive θ computable from Formula (5) is $\sqrt{\eta}$, that is, approximately 1×10^{-8} rad for IEEE arithmetic. Thus, an angle smaller than 1×10^{-8} rad cannot be detected by the application of Formula (5) (unless the angle is computed as zero). Conversely, the alternative Formula (6) can be expected to return values of η having small relative errors of at most a small multiple of η for all $\boldsymbol{a}$ and $\boldsymbol{b}$.

This example has consequences for the implementation of the procedures described in ISO 10360-6:2001 concerned with testing software for computing Gaussian best-fit geometric features to measured data. Such computations are used in industrial inspection and geometric tolerancing applications. The Standard requires that the unit direction vector used in the parametrization of such geometric features as planes, cylinders and cones and returned by test software is compared with a reference solution by computing the angle between the test and reference vectors. The Standard defines acceptance of the test vector if this angle is smaller than 1×10^{-8} rad. If Formula (5) were used to evaluate the angle, this acceptance criterion would never be satisfied no matter how close were the test and reference vectors, unless the angle were computed as zero. There would be no problem with undertaking the comparison if Formula (6) were used. The importance of using a stable formula is recognized in the Standard, which includes information to this effect. Software for this calculation constitutes an example of implementation C of 5.1.3.

NOTE The computation of θ using Formula (6) involves taking the difference $\boldsymbol{b} - \boldsymbol{a}$ between two almost identical floating-point vectors, which can very often be formed exactly in IEEE arithmetic. The other operations (division by 2, taking a vector norm, determining $\sin^{-1}$ of the resulting value, and multiplication by 2) are innocuous regarding rounding error and have relative rounding errors of order η.

6.3 Maintenance of software

6.3.1 General

Software requires maintenance over its lifetime. The Committee should state the body responsible for maintaining the software. It is recommended that information concerning software revision is made available on the appropriate website (see Note 1 to 3.5).

6.3.2 Accessibility

ISO should at any time have access to software used in developing the publication.

6.3.3 Durability

6.3.3.1 Normally, an ISO publication is considered for revision every five years. Therefore, the durability of the software used for the development and application of the ISO publication should have a ‘shelf life’ of at least five years. Executable programs should be suitable in this regard. The same applies for most commercial standard statistical packages. However, there may be problems when a macro is used in a newer version of a standard statistical package or a different or revised compiler is used. All macros should as a principle be checked when there is a change of this type. The Committee should consider responsibility for and maintenance of the software used in the development and application of an ISO publication.

6.3.3.2 A working group is disbanded when its work has been completed. The ISO publications it produced, and any associated software, might remain relevant and useful, in which case it is recommended that regarding software support the work be allocated to the Committee.

7 Legal implications

There may be legal implications concerning software relating to the following aspects that should be considered:

a) commercial software suppliers named in an ISO publication;

b) no recommendation should be made or implied;

c) limitation of liability applies to the provision of software and reference data;

d) suitable copyright statements should be included. It may be appropriate to issue a separate licence to permit copying of parts of an ISO publication, such as sample code or data given in examples, so that third parties can maintain a publicly available electronic copy. Copyright considerations should be related to ISO publication policies.

It is beyond the scope of this Technical Report to consider such aspects in more detail. It is the responsibility of the Committee, with advice from ISO Central Secretariat, to address such issues.

8 Illustration

8.1 General

8.1.1 This clause illustrates the use of this Technical Report in developing an ISO publication concerned with methods of calculation that require the generation of pseudo-random numbers from a rectangular distribution.

NOTE Some of the material in this clause is extracted with minor adaptation from GUM Supplement 1[16].

8.1.2 For this illustration, considerations that apply in support categories 2 and 5 (see 5.1.1) are given.

8.1.3 Subclauses 8.2.1 and 8.2.2 typify some of the content of the part of the ISO publication concerned with such pseudo-random number generation in each of these cases. In both cases that part of the ISO publication would include text such as contained in 8.1.4.

8.1.4 To avoid the possibility of invalid results from these calculations, it is necessary that the pseudo-random number generator used has good statistical properties. The test suite TestU01[17] carries out an extensive test of the statistical properties of any generator submitted to it.

8.2 Examples of particular software

8.2.1 Support category 2: Specific software provided by a third party

A recommended generator that has passed the suite of tests indicated in 8.1.4 is the Mersenne Twister generator[18]. Software implementations are available in many languages including C, C++, C#, Fortran 95, Java, Lisp, Mathematica, MATLAB, Microsoft Excel, Perl and Python.[19]

8.2.2 Support category 5: A step-by-step procedure in a form for implementation as software by users or a third party

8.2.2.1 General

The enhanced Wichmann-Hill generator (8.2.2.2) has variants for 32- and 64-bit computers, passes the suite of tests indicated in 8.1.4, and has the properties[21] that

a) it is compact,

b) it is straightforward to code in any programming language that supports floating-point arithmetic, and does not depend upon the bit manipulation as used by some generators,

c) the state (the amount of information preserved by the generator between calls to it), namely the parameters i_1, i_2, i_3 and i_4 in Table 1, is small and easy to handle, and

d) it can readily be used to provide multiple sequences needed for highly parallel applications, likely to be a feature of large-scale calculations.

8.2.2.2 Procedure for generating pseudo-random numbers from a rectangular distribution

Table 1 defines the enhanced Wichmann-Hill generator for generating pseudo-random numbers from R(0, 1), a rectangular distribution on the interval (0, 1), for a 32-bit computer.

Table 1 — Enhanced Wichmann-Hill generator for pseudo-random numbers from a rectangular distribution on the interval (0, 1) for 32-bit computers

Input-output parameter	**Description**
i_1, i_2, i_3 and i_4	Integer parameters used as input quantities and changed by the procedure. Set to integers between 1 and 2 147 483 647 before first call. Do not disturb between calls. Subsequent values are not usually of concern to the user. These parameters provide the basis by which the pseudo-random numbers are generated. They may be realized as global variables and thus not appear explicitly as parameters
Constant	**Description**
a, b, c, d Do not disturb between calls	Vectors of integer constants of dimension 1×4, where $a = (a_1, a_2, a_3, a_4)$, etc., given by $a =$ (11 600, 47 003, 23 000, 33 000), $b =$ (185 127, 45 688, 93 368, 65 075), $c =$ (10 379, 10 479, 19 423, 8 123), $d =$ 2 147 483 123 × (1, 1, 1, 1) + (456, 420, 300, 0).
Output parameter	**Description**
r	Pseudo-random number drawn from R(0, 1)
	Computation
	For $j = 1, \ldots, 4$: i) Form $i_j = a_j \times (i_j \bmod b_j) - c_j \times \lfloor i_j / b_j \rfloor$ ii) If $i_j < 0$, replace i_j by $i_j + d_j$ Form $w = i_1/d_1 + i_2/d_2 + i_3/d_3 + i_4/d_4$ Form $r = w - \lfloor w \rfloor$
NOTE 1 $r = \lfloor w \rfloor$ denotes the largest integer no greater than w. NOTE 2 i_j mod b_j denotes the remainder on division of i_j by b_j.	

Bibliography

[1] IEEE 754-2008, *Standard for binary floating-point arithmetic*

[2] IEEE 854-1987, *A radix-independent standard for floating-point arithmetic*

[3] ISO 10360-6:2001, *Geometrical product specifications (GPS) — Acceptance test and reverification test for coordinate measuring machines (CMM) — Part 6: Computation of Gaussian associated features*

[4] ISO 13528:2005, *Statistical methods for use in proficiency testing by interlaboratory comparisons*

[5] ISO 19115:2003, *Geographic information — Metadata*

[6] ISO 80000-1:2009, *Quantities and units — General*

[7] ISO 9126-1:2001, *Software engineering — Product quality — Part 1: Quality model*

[8] ISO/IEC 18019:2004, *Software and system engineering — Guidelines for the design and preparation of user documentation for application software*

[9] ISO/TR 12845:2010, *Selected illustrations of fractional factorial screening experiments*

[10] ISO/TR 14468:2010, *Selected illustrations of attribute agreement analysis*

[11] ISO/TR 29901:2007, *Selected illustrations of full factorial experiments with four factors*

[12] ISO/TS 28037:2010, *Determination and use of straight-line calibration functions*

[13] Butler, B.P., Cox, M.G., Ellison, S.L.R. and Hardcastle, W.A., Eds. *Statistics Software Qualification: Reference Data Sets*. Royal Society of Chemistry, Cambridge, 1996

[14] Cook. H.R., Cox, M.G., Dainton, M.P. and Harris, P.M. *A methodology for testing spreadsheets and other packages used in metrology. Tech. Rep. CMSC 25/99*. National Physical Laboratory, Teddington, UK, 1999

[15] Haneef, N.J. Software documentation and readability: A proposed process improvement. *ACM SIGSOFT Software Engineering Notes*. 1998, **23**, pp. 75–77

[16] ISO/IEC Guide 98-3:2008, Suppl. 1:2008. *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) — Supplement 1: Propagation of distributions using a Monte Carlo method*

[17] L'Ecuyer, P. and Simard, R. TestU01: A software library in ANSI C for empirical testing of random number generators. http://www.iro.umontreal.ca/~simardr/testu01/tu01.html

[18] Matsumoto, M. and Nishimura, T. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. Model. Comput. Simul.* 1998, **8**, pp. 3–30

[19] Mersenne Twister, Wikipedia. http://en.wikipedia.org/wiki/Mersenne_twister/

[20] Statistical Reference Datasets (StRD), NIST. http://www.nist.gov/itl/sed/gsg/strd.cfm

[21] Wichmann, B. A. and Hill, I. D. Generating good pseudo-random numbers. *Comput. Stat. Data Anal.* 2006, **51**, pp. 1614–1622

This page deliberately left blank

British Standards Institution (BSI)

BSI is the national body responsible for preparing British Standards and other standards-related publications, information and services.

BSI is incorporated by Royal Charter. British Standards and other standardization products are published by BSI Standards Limited.

About us

We bring together business, industry, government, consumers, innovators and others to shape their combined experience and expertise into standards -based solutions.

The knowledge embodied in our standards has been carefully assembled in a dependable format and refined through our open consultation process. Organizations of all sizes and across all sectors choose standards to help them achieve their goals.

Information on standards

We can provide you with the knowledge that your organization needs to succeed. Find out more about British Standards by visiting our website at bsigroup.com/standards or contacting our Customer Services team or Knowledge Centre.

Buying standards

You can buy and download PDF versions of BSI publications, including British and adopted European and international standards, through our website at bsigroup.com/shop, where hard copies can also be purchased.

If you need international and foreign standards from other Standards Development Organizations, hard copies can be ordered from our Customer Services team.

Subscriptions

Our range of subscription services are designed to make using standards easier for you. For further information on our subscription products go to bsigroup.com/subscriptions.

With **British Standards Online (BSOL)** you'll have instant access to over 55,000 British and adopted European and international standards from your desktop. It's available 24/7 and is refreshed daily so you'll always be up to date.

You can keep in touch with standards developments and receive substantial discounts on the purchase price of standards, both in single copy and subscription format, by becoming a **BSI Subscribing Member**.

PLUS is an updating service exclusive to BSI Subscribing Members. You will automatically receive the latest hard copy of your standards when they're revised or replaced.

To find out more about becoming a BSI Subscribing Member and the benefits of membership, please visit bsigroup.com/shop.

With a **Multi-User Network Licence (MUNL)** you are able to host standards publications on your intranet. Licences can cover as few or as many users as you wish. With updates supplied as soon as they're available, you can be sure your documentation is current. For further information, email bsmusales@bsigroup.com.

Revisions

Our British Standards and other publications are updated by amendment or revision.

We continually improve the quality of our products and services to benefit your business. If you find an inaccuracy or ambiguity within a British Standard or other BSI publication please inform the Knowledge Centre.

Copyright

Useful Contacts:

Customer Services
Tel: +44 845 086 9001
Email (orders): orders@bsigroup.com
Email (enquiries): cservices@bsigroup.com

Subscriptions
Tel: +44 845 086 9001
Email: subscriptions@bsigroup.com

Knowledge Centre
Tel: +44 20 8996 7004
Email: knowledgecentre@bsigroup.com

Copyright & Licensing
Tel: +44 20 8996 7070
Email: copyright@bsigroup.com

BSI Group Headquarters

389 Chiswick High Road London W4 4AL UK

bsi.

...making excellence a habit.™