TECHNICAL SPECIFICATION

ISO/TS 21748

First edition
2004-03-15

Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation

Lignes directrices relatives à l'utilisation d'estimations de la répétabilité, de la reproductibilité et de la justesse dans l'évaluation de l'incertitude de mesure

Reference number
ISO/TS 21748:2004(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

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 other circumstances, particularly when there is an urgent market requirement for such documents, a technical committee may decide to publish other types of normative document:

— an ISO Publicly Available Specification (ISO/PAS) represents an agreement between technical experts in an ISO working group and is accepted for publication if it is approved by more than 50 % of the members of the parent committee casting a vote;

— an ISO Technical Specification (ISO/TS) represents an agreement between the members of a technical committee and is accepted for publication if it is approved by 2/3 of the members of the committee casting a vote.

An ISO/PAS or ISO/TS is reviewed after three years in order to decide whether it will be confirmed for a further three years, revised to become an International Standard, or withdrawn. If the ISO/PAS or ISO/TS is confirmed, it is reviewed again after a further three years, at which time it must either be transformed into an International Standard or be withdrawn.

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/TS 21748 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

Introduction

Knowledge of the uncertainty associated with measurement results is essential to the interpretation of the results. Without quantitative assessments of uncertainty, it is impossible to decide whether observed differences between results reflect more than experimental variability, whether test items comply with specifications, or whether laws based on limits have been broken. Without information on uncertainty, there is a risk of misinterpretation of results. Incorrect decisions taken on such a basis may result in unnecessary expenditure in industry, incorrect prosecution in law, or adverse health or social consequences.

Laboratories operating under ISO 17025 accreditation and related systems are accordingly required to evaluate measurement uncertainty for measurement and test results and report the uncertainty where relevant. The *Guide to the expression of uncertainty in measurement* (GUM), published by ISO, is a widely adopted standard approach. However, it applies to situations where a model of the measurement process is available. A very wide range of standard test methods is, however, subjected to collaborative study in accordance with ISO 5725-2:1994. This Technical Specification provides an appropriate and economic methodology for estimating uncertainty associated with the results of these methods, which complies fully with the relevant principles of the GUM, whilst taking account of method performance data obtained by collaborative study.

The general approach used in this Technical Specification requires that

— estimates of the repeatability, reproducibility and trueness of the method in use, obtained by collaborative study as described in ISO 5725-2:1994, be available from published information about the test method in use. These provide estimates of the intra- and inter-laboratory components of variance, together with an estimate of uncertainty associated with the trueness of the method;

— the laboratory confirm that its implementation of the test method is consistent with the established performance of the test method by checking its own bias and precision. This confirms that the published data are applicable to the results obtained by the laboratory;

— any influences on the measurement results that were not adequately covered by the collaborative study be identified and the variance associated with the results that could arise from these effects be quantified.

An uncertainty estimate is made by combining the relevant variance estimates in the manner prescribed by the GUM.

The dispersion of results obtained in a collaborative study is often also usefully compared with measurement uncertainty estimates obtained using GUM procedures as a test of full understanding of the method. Such comparisons will be more effective given a consistent methodology for estimating the same parameter using collaborative study data.

Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation

1 Scope

The Technical Specification gives guidance for

— evaluation of measurement uncertainties using data obtained from studies conducted in accordance with ISO 5725-2:1994;

— comparison of collaborative study results with measurement uncertainty (MU) obtained using formal principles of uncertainty propagation (see Clause 14).

ISO 5725-3:1994 provides additional models for studies of intermediate precision. However, while the same general approach may be applied to the use of such extended models, uncertainty evaluation using these models is not incorporated in the present Technical Specification.

This Technical Specification is applicable in all measurement and test fields where an uncertainty associated with a result has to be determined.

This Technical Specification does not describe the application of repeatability data in the absence of reproducibility data.

This Technical Specification assumes that recognized, non-negligible systematic effects are corrected, either by applying a numerical correction as part of the method of measurement, or by investigation and removal of the cause of the effect.

The recommendations in this Technical Specification are primarily for guidance. It is recognized that while the recommendations presented do form a valid approach to the evaluation of uncertainty for many purposes, it is also possible to adopt other suitable approaches.

In general, references to measurement results, methods and processes in this Technical Specification are normally understood to apply also to testing results, methods and processes.

2 Normative references

The following referenced documents are indispensable for the application 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 3534-1, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*

ISO 5725-3:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate measures of the precision of a standard measurement method*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply. In addition, reference is made to "intermediate precision conditions", which are discussed in detail in ISO 5725-3:1994.

3.1
bias
difference between the expectation of the test results and an accepted reference value

NOTE Bias is the total systematic error as contrasted to random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.

[ISO 3534-1]

3.2
combined standard uncertainty
$u(y)$
standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in these quantities

[GUM]

3.3
coverage factor
k
numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

NOTE A coverage factor, k, is typically in the range 2 to 3.

[GUM]

3.4
expanded uncertainty
U
quantity defining an interval about a result of a measurement expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand

NOTE 1 The fraction may be regarded as the coverage probability or level of confidence of the interval.

NOTE 2 To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterised by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions can be justified.

NOTE 3 Expanded uncertainty is termed overall uncertainty in Recommendation INC-1 (1980), paragraph 5.

[GUM]

3.5
precision
closeness of agreement between independent test results obtained under stipulated conditions

NOTE 1 Precision depends upon the distribution of random errors and does not relate to the true value or the specified value.

NOTE 2 The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Less precision is reflected by a higher standard deviation.

NOTE 3 "Independent test results" means results obtained in a manner not influenced by any previous result on the same or similar test object. Quantitative measures of precision depend critically on the stipulated conditions. Repeatability and reproducibility conditions are particular examples of extreme stipulated conditions.

[ISO 3534-1]

3.6
repeatability
precision under repeatability conditions, i.e. conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time

[ISO 3534-1]

3.7
repeatability standard deviation
standard deviation of test results obtained under repeatability conditions

NOTE This is a measure of dispersion of the distribution of test results under repeatability conditions. Similarly "repeatability variance" and "repeatability coefficient of variation" can be defined and used as measures of the dispersion of test results under repeatability conditions.

[ISO 3534-1]

3.8
reproducibility
precision under reproducibility conditions, i.e. conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment

NOTE A valid statement of reproducibility requires specification of the conditions changed. Reproducibility may be expressed quantitatively in terms of the dispersion of the results.

[ISO 3534-1]

3.9
reproducibility standard deviation
standard deviation of test results obtained under reproducibility conditions

NOTE This is a measure of dispersion of the distribution of test results under reproducibility conditions. Similarly "reproducibility variance" and "reproducibility coefficient of variation" could be defined and used as measures of the dispersion of test results under reproducibility conditions.

[ISO 3534-1]

3.10
standard uncertainty
$u(x_i)$
uncertainty of the result of a measurement expressed as a standard deviation

[GUM]

3.11
trueness
closeness of agreement between the average value obtained from a large set of test results and an accepted reference value

NOTE The measure of trueness is normally expressed in terms of bias. The reference to trueness as “accuracy of the mean” is not generally recommended.

[ISO 3534-1]

3.12
uncertainty
⟨measurement⟩ parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand

NOTE 1 The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

NOTE 2 Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of a series of measurements and can be characterized by experimental standard deviations. Other components, which also can be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.

NOTE 3 It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects such as components associated with corrections and reference standards, contribute to the dispersion.

[GUM]

3.13
uncertainty budget
list of sources of uncertainty and their associated standard uncertainties, compiled with a view to evaluating a combined standard uncertainty associated with a measurement result

NOTE The list often includes additional information such as sensitivity coefficients (rate of change of result with change in a quantity affecting the result), degrees of freedom for each standard uncertainty, and an identification of the means of evaluating each standard uncertainty in terms of a Type A or Type B evaluation.

4 Symbols

a coefficient indicating an intercept in the empirical relationship $\hat{s}_R = a + bm$

B laboratory component of bias

b coefficient indicating a slope in the empirical relationship $\hat{s}_R = a + bm$

c coefficient in the empirical relationship $\hat{s}_R = cm^d$

c_i sensitivity coefficient $\partial y / \partial x_i$

d coefficient indicating an exponent in the empirical relationship $\hat{s}_R = cm^d$

e random residual error

e_r random residual error under repeatability conditions

k numerical factor used as a multiplier of the combined standard uncertainty u in order to obtain an expanded uncertainty U

l laboratory number

m mean value of the measurements

N number of contributions included in combined uncertainty calculations

n' number of contributions incorporated in combined uncertainty calculations in addition to collaborative study data

n_l number of replicates at one level by laboratory l

n_r number of replicate measurements

p number of laboratories

Q number of test items from a larger batch

q number of assigned values by consensus during a collaborative study

r_{ij} correlation coefficient between x_i and x_j, in the range -1 to $+1$

s_b between-group component of variance expressed as a standard deviation

s_{b}^2 between-group component of variance

s_{D} estimated, or experimental, standard deviation of results obtained by repeated measurement on a reference material used for checking control of bias

s_i repeatability standard deviation with ν_i degrees of freedom

s_{inh} uncertainty associated with the inhomogeneity of the sample

s_{inh}^2 component of variance associated with the inhomogeneity of the sample

s_{L} experimental or estimated inter-laboratory standard deviation

$\hat{s}_{\mathrm{L}}$ adjusted uncertainty associated with B where the contribution is dependent on the response

s_{L}^2 the estimated variance of B

s_r intra-laboratory standard deviation

$\hat{s}_r$ adjusted estimate of inter-laboratory standard deviation, where the contribution is dependent on the response

s_r^2 estimated variance of e_r

s_R estimated reproducibility standard deviation

s'_R adjusted estimate of the reproducibility standard deviation

$\hat{s}_R$ adjusted reproducibility standard deviation calculated from an empirical model, where the contributions are dependent on the response

s_{w} intra-laboratory standard deviation derived from replicates or other repeatability studies

s_{w}^2 intra-group component of variance (often an intra-laboratory component of variance)

$s(\Delta_y)$ laboratory standard deviation of differences during a comparison of a routine method with a definitive method

x_i value of the ith input value in the determination of a result

x'_i deviation of the ith input value from the nominal value of x

x_j the jth input value in the determination of a result

$u(\hat{\delta})$ uncertainty associated with δ due to the uncertainty of estimating δ by measuring a reference measurement standard or reference material with certified value $\hat{\mu}$

$u(\hat{\mu})$ uncertainty associated with the certified value $\hat{\mu}$

$u(y)$ combined standard uncertainty associated with y where $u(y)=\sqrt{\sum_{i=1,n} c_i^2 u^2(x_i)}$

$u(Y)$ combined uncertainty for the result $Y = f(y_1, y_2, ...)$ where $u(Y)=\sqrt{\sum_i \left[c_i u(y_i)\right]^2}$

$u^2(y)$ combined standard uncertainty associated with y, expressed as a variance

u_{inh} uncertainty associated with sample inhomogeneity

U expanded uncertainty, equal to k times the standard uncertainty u

$U(y)$ expanded uncertainty in y where $U(y) = ku(y)$, where k is a coverage factor

y_i result for test item i from the definitive method during a comparison of methods

$\hat{y}_i$ result for test item i from the routine test method during a comparison of methods

y_0 assigned value for proficiency testing

Δ laboratory bias

Δ_l estimate of bias of laboratory l, equal to the laboratory mean, m, minus the certified value, $\hat{\mu}$

$\bar{\Delta}_y$ mean laboratory bias during a comparison of a routine method with a definitive method

δ bias intrinsic to the measurement method in use

$\hat{\delta}$ estimated or measured bias

μ unknown expectation of the ideal result

$\hat{\mu}$ certified value of a reference material

σ_0 standard deviation for proficiency testing

σ_D true value of the standard deviation of results obtained by repeated measurement on a reference material used for checking control of bias

σ_L inter-laboratory standard deviation; standard deviation of B

σ_L^2 variance of B; inter-laboratory variance

σ_r intra-laboratory standard deviation; standard deviation of e_r

σ_r^2 variance of e_r; intra-laboratory variance

σ_w within-group standard deviation

σ_{w0} standard deviation required for adequate performance (ISO Guide 33)

ν_{eff} effective degrees of freedom for the standard deviation of, or uncertainty associated with input value x_i