SLOVENSKI STANDARD
SIST ISO 5725-3:2024

01-september-2024

Točnost (pravilnost in natančnost) merilnih metod in rezultatov – 3. del : Vmesne mere natančnosti in alternativni pristopi za primerjalne študije

Accuracy (trueness and precision) of measurement methods and results — Part 3: Intermediate precision and alternative designs for collaborative studies

Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 3: Fidélité intermédiaire et plans alternatifs pour les études collaboratives

Ta slovenski standard je istoveten z: **ISO 5725-3:2023**

ICS:

03.120.30	Uporaba statističnih metod	Application of statistical methods
17.020	Meroslovje in merjenje na splošno	Metrology and measurement in general

SIST ISO 5725-3:2024 **en,fr,de**

INTERNATIONAL STANDARD

ISO 5725-3

Second edition
2023-06

Accuracy (trueness and precision) of measurement methods and results —

Part 3: Intermediate precision and alternative designs for collaborative studies

Exactitude (justesse et fidélité) des résultats et méthodes de mesure —

Partie 3: Fidélité intermédiaire et plans alternatifs pour les études collaboratives

Reference number
ISO 5725-3:2023(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: 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.

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 document 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of 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 www.iso.org/iso/foreword.html.

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

This second edition cancels and replaces the first edition (ISO 5725-3:1994), which has been technically revised. It also incorporates the Technical Corrigendum ISO 5725-3:1994/Cor.1:2001.

The main changes are as follows:

— Several additional experimental designs have been added to this version compared to the previous version, some of them from ISO 5725-5. These are orthogonal array designs, split level designs, designs for heterogeneous sample material as well as designs across levels.

— Furthermore, the standard was supplemented by considerations on the selection of factors and modelling of the factorial effects, as well as by a section in which the reliability of the various interlaboratory test parameters (mean and precision parameters) are considered.

A list of all parts in the ISO 5725 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

0.1 ISO 5725 uses two terms “trueness” and "precision” to describe the accuracy of a measurement method. “Trueness” refers to the degree of agreement between the average value of a large number of test results and the true or accepted reference value. “Precision” refers to the degree of agreement between test results.

0.2 General consideration of these quantities is given in ISO 5725-1 and is not repeated here. It is stressed that ISO 5725-1 provides underlying definitions and general principles should be read in conjunction with all other parts of ISO 5725.

0.3 Many different factors (apart from test material heterogeneity) may contribute to the variability of results from a measurement method, including:

a) the laboratory;

b) the operator;

c) the equipment used;

d) the calibration of the equipment;

e) the batch of a reagent;

f) the time elapsed between measurements;

g) environment (temperature, humidity, air pollution, etc.);

h) other factors.

0.4 Two conditions of precision, termed repeatability and reproducibility conditions, have been found necessary and, for many practical cases, useful for describing the variability of a measurement method. Under repeatability conditions, none of the factors a) to h) in 0.3 are considered to vary, while under reproducibility conditions, all of the factors are considered to vary and contribute to the variability of the test results. Thus, repeatability and reproducibility conditions are the two extremes of precision, the first describing the minimum and the second the maximum variability in results. Intermediate conditions between these two extreme conditions of precision are also conceivable, when one or more of the factors listed in b) to g) are allowed to vary.

To illustrate the need for including a consideration of intermediate conditions in method validation, consider the operation of a present-day laboratory connected with a production plant involving, for example, a three-shift working system where measurements are made by different operators on different equipment. Operators and equipment are then some of the factors that contribute to the variability in the test results.

The standard deviation of test results obtained under repeatability conditions is generally less than that obtained under intermediate precision conditions. Generally, in chemical analysis, the standard deviation under intermediate precision conditions may be two or three times larger than that under repeatability conditions. It should not, of course, exceed the reproducibility standard deviation.

As an example, in the determination of copper in copper ore, a collaborative study among 35 laboratories revealed that the standard deviation under intermediate precision conditions (different times) was 1,5 times larger than that under repeatability conditions, both for the electrolytic gravimetry and $Na_2S_20_3$ titration methods.

0.5 This document focuses on intermediate precision and alternative designs for collaborative studies of a measurement method. Apart from the determination of intermediate precision measures, the aims of these alternative designs include reducing the number of required measurements, increasing the reliability of the estimates for precision and overall mean and taking into account test material heterogeneity.

Indeed, a t-factor fully-nested experiment with two levels per factor (inside each laboratory, there are $t-1$ factors) and two replicates per setting requires $2 \cdot 2^{t-1}$ test results from each laboratory, which can be an excessive requirement on the laboratories. For this reason, in the previous version of ISO 5725-3, the staggered nested design is also discussed. While the estimation of the precision parameters is more complex and subject to greater uncertainty in a staggered nested design, the workload is reduced. This document offers alternative strategies to reduce the workload without compromising the reliability of the precision estimates.

As far as the special designs for sample heterogeneity are concerned, they were discussed in the previous version of ISO 5725-5. However, it is convenient to have one part of this standard dedicated to the question of the design of experiments.

0.6 The repeatability precision as determined in accordance with ISO 5725-2 is computed as a mean across participating laboratories. Whether it can be used for quality control purposes depends on whether the repeatability standard deviation can be considered to remain constant across laboratories. For this reason, it is important to obtain information on how the repeatability standard deviation varies within and between the laboratories under different conditions.

0.7 In many collaborative studies, the between-laboratory variability is large in comparison to the repeatability, and it would be useful to a) decompose it into several different precision components, b) reduce, if possible, some sources of variability which are due to the intermediate precision conditions. This can be done by identifying factors (e.g. time, calibration, operator or equipment) which contribute to the variability under intermediate precision conditions of measurement, by quantifying the corresponding variability components and, wherever achievable, decreasing their contribution. In this manner, the intermediate precision component of the overall variance is enlarged while the between-laboratory component of the overall variance is reduced. Only random effects are considered: it is only reasonable to model a factor as a fixed effect after a method or calibration optimization study has been conducted. In this standard, different relationships between factors are taken into account, e.g. whether a particular factor is subsumed under another factor or not.

0.8 Estimates for precision and overall mean are subject to random variability. Accordingly, it is important to determine the uncertainty associated with each estimate, and to understand the relationships between this uncertainty, the number of participants and the design. Once these relationships are understood, it becomes possible to make much more informed decisions concerning the number of participants and the experimental design.

0.9 Provided different factorial effects do contribute to the variability, determining the respective precision components may make it possible to reduce the required number of participating laboratories, since the between-laboratory variability can be expected to be less dominant. However, it is highly recommended to have a reasonable number of participating laboratories in order to ensure a realistic assessment of the overall method variability obtained under routine conditions of operation.

0.10 In the uniform-level design according to part 2 of this standard, there is a risk that an operator will allow the result of a measurement on one sample to influence the result of a subsequent measurement on another sample of the same material, causing the estimates of the repeatability and reproducibility standard deviations to be biased. When this risk is considered to be serious, the split-level design described in this document may be preferred as it reduces this risk. Care should be taken that the two materials used at a particular level of the experiment are sufficiently similar to ensure that the same precision measures can be expected (in other words: the question arises whether the precision component associated with a particular factor remains unchanged across a range of similar matrices).

0.11 The experimental design presented in ISO 5725-2 requires the preparation of a number of identical samples of the material for use in the experiment. With heterogeneous materials this may not be possible, so that the use of the basic method then gives estimates of the reproducibility standard deviation that are inflated by the variation between the samples. The design for a heterogeneous material given in this document yields information about the variability between samples which is not obtainable from the basic method; it may be used to calculate an estimate of reproducibility from which the between-sample variation has been removed.

Accuracy (trueness and precision) of measurement methods and results —

Part 3: Intermediate precision and alternative designs for collaborative studies

1 Scope

This document provides

a) a discussion of alternative experimental designs for the determination of trueness and precision measures including reproducibility, repeatability and selected measures of intermediate precision of a standard measurement method, including a review of the circumstances in which their use is necessary or beneficial, and guidance as to the interpretation and application of the resulting estimates, and

b) worked examples including specific designs and computations.

Each of the alternative designs discussed in this document is intended to address one (or several) of the following issues:

a) a discussion of the implications of the definitions of intermediate precision measures;

b) a guidance on the interpretation and application of the estimates of intermediate precision measures in practical situations;

c) determining reproducibility, repeatability and selected measures of intermediate precision;

d) improved[1)] determination of reproducibility and other measures of precision;

e) improving the estimate of the sample mean;

f) determining the range of in-house repeatability standard deviations;

g) determining other precision components such as operator variability;

h) determining the level of reliability of precision estimates;

i) reducing the minimum number of participating laboratories by optimizing the reliability of precision estimates;

j) avoiding distorted estimations of repeatability (split-level designs);

k) avoiding distorted estimations of reproducibility (taking the heterogeneity of the material into consideration).

Often, the performance of the method whose precision is being evaluated in a collaborative study will have previously been assessed in a single-laboratory validation study conducted by the laboratory which developed it. Relevant factors for the determination of intermediary precision will have been identified in this prior single-laboratory study.

1) Allowing a reduction in the number of laboratories.

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 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

ISO 3534-2, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

ISO 5725-1, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*

ISO Guide 33, *Reference materials — Good practice in using reference materials*

ISO Guide 35, *Reference materials — Guidance for characterization and assessment of homogeneity and stability*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1, ISO 3534-2 and ISO 5725-1 and the following 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 https://www.electropedia.org/

3.1
block
group of *settings* (3.7) conducted in parallel or within a short time interval, and with the same samples

EXAMPLE Two settings:

Operator 1 + Calibration 1 + Equipment 1 + Batch 1

and

Operator 1 + Calibration 2 + Equipment 2 + Batch 1

Note 1 to entry: This definition is more specific than the general definition given in ISO 3534-3:2013, 3.1.25, where block is defined as a collection of experimental units.

3.2
factor
feature under examination as a potential source of variation

EXAMPLE Operator, calibration, equipment, day, reagent batch, storage temperature, shaker orbit, shaker frequency.

Note 1 to entry: Strictly speaking, the factor laboratory is a factor just like any other. However, since the ISO 5725 standard focuses on method validation by means of interlaboratory studies, the factor laboratory can be considered to have a somewhat privileged role. The following characteristics distinguish it from other factors:

— The factor laboratory is indispensable: For each measurement, the name of the particular laboratory where it was performed will *always* be provided in a collaborative study.

— The factor laboratory will almost always have more levels than other factors.

It should also be noted that categories such as measurand, sample/matrix and level may also be considered to be factors. However, in collaborative studies, they are often not taken into account as such in the factorial design. The reason is that, for these factors, one is interested in a separate statistical analysis for each separate factor level. In other words, one is interested in obtaining separate precision measures for each particular measurand or concentration level, not across measurands or concentration levels. However, in cases where it is required to quantify precision across, say, matrices, then the factor sample/matrix should also be included in the design. Accordingly, in this document, designs are discussed to be applied for a particular measurand or concentration level by different laboratories all applying the same measurement procedure.

[SOURCE: ISO 3534-3:2013, 3.1.5, modified — Note 1 to entry was modified and Note 2 to entry was deleted.]

3.3
factor level
setting (3.7), value or assignment of a *factor* (3.2)

EXAMPLE Operator 1, Operator 2

Note 1 to entry: In many designs, the majority of factors will be varied across two levels.

3.4
fully-nested design
nested design, where there is a nesting hierarchy for every pair of *factors* (3.2)

EXAMPLE There are 2 operators in each laboratory, and each operator performs 2 calibrations, i.e., the study includes 2 operators and 4 calibrations for each laboratory.

3.5
partially-nested design
nested design where one *factor* (3.2) (the factor laboratory) is ranked higher than all other factors (i.e., all other factors are nested within the factor laboratory), and there is at least one factor pair without a nesting hierarchy

EXAMPLE There are 2 operators and 2 instruments in each laboratory, and each operator performs measurements on 2 instruments, i.e., the study includes 2 operators and 2 instruments for each laboratory.

3.6
run
actual measurement carried out for a particular *setting* (3.7) and for a particular laboratory

EXAMPLE Operator 1 + Equipment 1 + Batch 1 + Day 1 carried out in laboratory 1

Note 1 to entry: This definition is more specific than the general definition given in ISO 3534-3 (3.1.13), where run is defined as specific settings of every factor used on a particular experimental unit.

Note 2 to entry: "Identical" runs are called *replicates*, whereby "identical" means that the different time points are close enough to each other to allow for the results to be considered as obtained under repeatability conditions.

3.7
setting
combination of *factor levels* (3.3), for all *factors* (3.2) except the factor laboratory

EXAMPLE Operator 1 + Equipment 1 + Batch 1 + Day 1.

4 Symbols

B	Component in a test result representing the deviation of a laboratory from the general average (laboratory component of bias)

B_0	Component of B representing all factors that do not vary under intermediate precision conditions – laboratory bias *per se*
$B_{(1)}, B_{(2)}$, etc.	Components of B representing factors that vary under intermediate precision conditions
e	Component representing the random error occurring in every test result, corresponding to the analytical, repeatability, model or residual error
m	Overall mean of the measurand or test property for a particular matrix; level
$\hat{m}$	Estimate of the overall mean
n	Number of replicate test results obtained in one laboratory at one level for one setting
p	Number of laboratories participating in the collaborative study
q	Number of levels of the test property in the collaborative study
σ_w	Within-laboratory standard deviation of the residual term e
σ_r	Repeatability standard deviation
σ_R	Reproducibility standard deviation
σ_0	Standard deviation corresponding to factor B_0
$\sigma_{(1)}$	Standard deviation corresponding to factor $B_{(1)}$
$\sigma_{(2)}$	Standard deviation corresponding to factor $B_{(2)}$
σ_A	Standard deviation corresponding to factor A
$\sigma_{Interaction}$	Standard deviation corresponding to the interaction of two factors
σ_{AB}	Standard deviation corresponding to the interaction of the two factors A and B
s	Estimate of a standard deviation
se	Standard error
$Var(X)$	Variance of X
w	Range of a set of test results
y	Test result
$\bar{X}$	Mean of X
$\lvert X \rvert$	Absolute value of X

5 General requirements

In order to ensure that measurements are carried out in the same way, the measurement method shall have been standardized. All measurements obtained in the framework of an experiment within a specific laboratory or of a collaborative study shall be carried out according to that standard.

NOTE The terms collaborative experiment, collaborative trial and interlaboratory experiment are used interchangeably to denote a collaborative study conducted in order to characterize and/or assess the performance of a measurement method.