INTERNATIONAL STANDARD

ISO 5725-3

First edition
1994-12-15

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

Part 3: Intermediate measures of the precision of a standard measurement method

Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 3: Mesures intermédiaires de la fidélité d'une méthode de mesure normalisée

Reference number
ISO 5725-3:1994(E)

Contents

International Organization for Standardization
Case Postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

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.

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.

International Standard ISO 5725-3 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

ISO 5725 consists of the following parts, under the general title *Accuracy (trueness and precision) of measurement methods and results*:

— *Part 1: General principles and definitions*

— *Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*

— *Part 3: Intermediate measures of the precision of a standard measurement method*

— *Part 4: Basic methods for the determination of the trueness of a standard measurement method*

— *Part 5: Alternative methods for the determination of the precision of a standard measurement method*

— *Part 6: Use in practice of accuracy values*

Parts 1 to 6 of ISO 5725 together cancel and replace ISO 5725:1986, which has been extended to cover trueness (in addition to precision) and intermediate precision conditions (in addition to repeatability conditions and reproducibility conditions).

Annexes A, B and C form an integral part of this part of ISO 5725. Annexes D and E are for information only.

Introduction

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

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

0.3 Many different factors (apart from variations between supposedly identical specimens) may contribute to the variability of results from a measurement method, including:

a) the operator;

b) the equipment used;

c) the calibration of the equipment;

d) the environment (temperature, humidity, air pollution, etc.);

e) the batch of a reagent;

f) the time elapsed between measurements.

The variability between measurements performed by different operators and/or with different equipment will usually be greater than the variability between measurements carried out within a short interval of time by a single operator using the same equipment.

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, factors a) to f) in 0.3 are considered constants and do not contribute to the variability, while under reproducibility conditions they vary and do 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 factors

a) to f) are allowed to vary, and are used in certain specified circumstances.

Precision is normally expressed in terms of standard deviations.

0.5 This part of ISO 5725 focuses on intermediate precision measures of a measurement method. Such measures are called intermediate as their magnitude lies between the two extreme measures of the precision of a measurement method: repeatability and reproducibility standard deviations.

To illustrate the need for such intermediate precision measures, 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. These factors need to be taken into account when assessing the precision of the measurement method.

0.6 The intermediate precision measures defined in this part of ISO 5725 are primarily useful when their estimation is part of a procedure that aims at developing, standardizing, or controlling a measurement method within a laboratory. These measures can also be estimated in a specially designed interlaboratory study, but their interpretation and application then requires caution for reasons explained in 1.3 and 9.1.

0.7 The four factors most likely to influence the precision of a measurement method are the following.

a) **Time:** whether the time interval between successive measurements is short or long.

b) **Calibration:** whether the same equipment is or is not recalibrated between successive groups of measurements.

c) **Operator:** whether the same or different operators carry out the successive measurements.

d) **Equipment:** whether the same or different equipment (or the same or different batches of reagents) is used in the measurements.

0.8 It is, therefore, advantageous to introduce the following M-factor-different intermediate precision conditions ($M = 1$, 2, 3 or 4) to take account of changes in measurement conditions (time, calibration, operator and equipment) within a laboratory.

a) $M = 1$: only one of the four factors is different;

b) $M = 2$: two of the four factors are different;

c) $M = 3$: three of the four factors are different;

d) $M = 4$: all four factors are different.

Different intermediate precision conditions lead to different intermediate precision standard deviations denoted by $s_{I(\,)}$, where the specific conditions are listed within the parentheses. For example, $s_{I(TO)}$ is the inter-

mediate precision standard deviation with different times (T) and operators (O).

0.9 For measurements under intermediate precision conditions, one or more of the factors listed in 0.7 is or are different. Under repeatability conditions, those factors are assumed to be constant.

The standard deviation of test results obtained under repeatability conditions is generally less than that obtained under the conditions for intermediate precision conditions. Generally in chemical analysis, the standard deviation under intermediate precision conditions may be two or three times as large as 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 experiment among 35 laboratories revealed that the standard deviation under one-factor-different intermediate precision conditions (operator and equipment the same but time different) was 1,5 times larger than that under repeatability conditions, both for the electrolytic gravimetry and $Na_2S_2O_3$ titration methods.

This page intentionally left blank

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

Part 3:
Intermediate measures of the precision of a standard measurement method

1 Scope

1.1 This part of ISO 5725 specifies four intermediate precision measures due to changes in observation conditions (time, calibration, operator and equipment) within a laboratory. These intermediate measures can be established by an experiment within a specific laboratory or by an interlaboratory experiment.

Furthermore, this part of ISO 5725

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

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

c) does not provide any measure of the errors in estimating intermediate precision measures;

d) does not concern itself with determining the trueness of the measurement method itself, but does discuss the connections between trueness and measurement conditions.

1.2 This part of ISO 5725 is concerned exclusively with measurement methods which yield measurements on a continuous scale and give a single value as the test result, although the single value may be the outcome of a calculation from a set of observations.

1.3 The essence of the determination of these intermediate precision measures is that they measure the ability of the measurement method to repeat test results under the defined conditions.

1.4 The statistical methods developed in this part of ISO 5725 rely on the premise that one can pool information from "similar" measurement conditions to obtain more accurate information on the intermediate precision measures. This premise is a powerful one as long as what is claimed as "similar" is indeed "similar". But it is very difficult for this premise to hold when intermediate precision measures are estimated from an interlaboratory study. For example, controlling the effect of "time" or of "operator" across laboratories in such a way that they are "similar", so that pooling information from different laboratories makes sense, is very difficult. Thus, using results from interlaboratory studies on intermediate precision measures requires caution. Within-laboratory studies also rely on this premise, but in such studies it is more likely to be realistic, because the control and knowledge of the actual effect of a factor is then more within reach of the analyst.

1.5 There exist other techniques besides the ones described in this part of ISO 5725 to estimate and to verify intermediate precision measures within a lab-

oratory, for example, control charts (see ISO 5725-6). This part of ISO 5725 does not claim to describe the only approach to the estimation of intermediate precision measures within a specific laboratory.

NOTE 1 This part of ISO 5725 refers to designs of experiments such as nested designs. Some basic information is given in annexes B and C. Other references in this area are given in annex E.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 5725. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 5725 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms.*

ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions.*

ISO 5725-2:1994, *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.*

ISO Guide 33:1989, *Uses of certified reference materials.*

ISO Guide 35:1989, *Certification of reference materials — General and statistical principles.*

3 Definitions

For the purposes of this part of ISO 5725, the definitions given in ISO 3534-1 and ISO 5725-1 apply.

The symbols used in ISO 5725 are given in annex A.

4 General requirement

In order that the measurements are made in the same way, the measurement method shall have been standardized. All measurements forming part of an experiment within a specific laboratory or of an inter-laboratory experiment shall be carried out according to that standard.

5 Important factors

5.1 Four factors (time, calibration, operator and equipment) in the measurement conditions within a laboratory are considered to make the main contributions to the variability of measurements (see table 1).

5.2 "Measurements made at the same time" include those conducted in as short a time as feasible in order to minimize changes in conditions, such as environmental conditions, which cannot always be guaranteed constant. "Measurements made at different times", that is those carried out at long intervals of time, may include effects due to changes in environmental conditions.

Table 1 — Four important factors and their states

Factor	Measurement conditions within a laboratory	
	State 1 (same)	State 2 (different)
Time	Measurements made at the same time	Measurements made at different times
Calibration	No calibration between measurements	Calibration carried out between measurements
Operator	Same operator	Different operators
Equipment	Same equipment without recalibration	Different equipment

5.3 "Calibration" does not refer here to any calibration required as an integral part of obtaining a test result by the measurement method. It refers to the calibration process that takes place at regular intervals between groups of measurements within a laboratory.

5.4 In some operations, the "operator" may be, in fact, a team of operators, each of whom performs some specific part of the procedure. In such a case, the team should be regarded as the operator, and any change in membership or in the allotment of duties within the team should be regarded as providing a different "operator".

5.5 "Equipment" is often, in fact, sets of equipment, and any change in any significant component should be regarded as providing different equipment. As to what constitutes a significant component, common sense must prevail. A change of thermometer would be considered a significant component, but using a slightly different vessel to contain a water bath would be considered trivial. A change of a batch of a reagent should be considered a significant component. It can lead to different "equipment" or to a recalibration if such a change is followed by calibration.

5.6 Under repeatability conditions, all four factors are at state 1 of table 1. For intermediate precision conditions, one or more factors are at state 2 of table 1, and are specified as "precision conditions with M factor(s) different", where M is the number of factors at state 2. Under reproducibility conditions, results are obtained by different laboratories, so that not only are all four factors at state 2 but also there are additional effects due to the differences between laboratories in management and maintenance of the laboratories, general training levels of operators, and in stability and checking of test results, etc.

5.7 Under intermediate precision conditions with M factor(s) different, it is necessary to specify which factors are at state 2 of table 1 by means of suffixes, for example:

— time-different intermediate precision standard deviation, $s_{I(T)}$;

— calibration-different intermediate precision standard deviation, $s_{I(C)}$;

— operator-different intermediate precision standard deviation, $s_{I(O)}$;

— [time + operator]-different intermediate precision standard deviation, $s_{I(TO)}$;

— [time + operator + equipment]-different intermediate precision standard deviation, $s_{I(TOE)}$;

— and many others in a similar fashion.

6 Statistical model

6.1 Basic model

For estimating the accuracy (trueness and precision) of a measurement method, it is useful to assume that every test result, y, is the sum of three components:

$$y = m + B + e \quad \ldots (1)$$

where, for the particular material tested,

m is the general mean (expectation);

B is the laboratory component of bias under repeatability conditions;

e is the random error occurring in every measurement under repeatability conditions.

A discussion of each of these components, and of extensions of this basic model, follows.

6.2 General mean, m

6.2.1 The general mean, m, is the overall mean of the test results. The value of m obtained in a collaborative study (see ISO 5725-2) depends solely on the "true value" and the measurement method, and does not depend on the laboratory, equipment, operator or time by or at which any test result has been obtained. The general mean of the particular material measured is called the "level of the test"; for example, specimens of different purities of a chemical or different materials (e.g. different types of steel) will correspond to different levels.

In many situations, the concept of a true value μ holds good, such as the true concentration of a solution which is being titrated. The level m is not usually equal to the true value μ; the difference $(m - \mu)$ is called the "bias of the measurement method".

In some situations, the level of the test is exclusively defined by the measurement method, and the concept of an independent true value does not apply; for example, the Vicker's hardness of steel and the Micum indices of coke belong to this category. However, in general, the bias is denoted by δ ($\delta = 0$ where no true value exists), then the general mean m is

$$m = \mu + \delta \qquad \ldots (2)$$

NOTE 2 Discussion of the bias term δ and a description of trueness experiments are given in ISO 5725-4.

6.2.2 When examining the difference between test results obtained by the same measurement method, the bias of the measurement method may have no influence and can be ignored, unless it is a function of the level of the test. When comparing test results with a value specified in a contract, or a standard value where the contract or specification refers to the true value μ and not to the level of the test m, or when comparing test results obtained using different measurement methods, the bias of the measurement method must be taken into account.

6.3 Term B

6.3.1 B is a term representing the deviation of a laboratory, for one or more reasons, from m, irrespective of the random error e occurring in every test result. Under repeatability conditions in one laboratory, B is considered constant and is called the "laboratory component of bias".

6.3.2 However, when using a measurement method routinely, it is apparent that embodied within an overall value for B are a large number of effects which are due, for example, to changes in the operator, the equipment used, the calibration of the equipment, and the environment (temperature, humidity, air pollution, etc.). The statistical model [equation (1)] can then be rewritten in the form:

$$y = m + B_0 + B_{(1)} + B_{(2)} + \ldots + e \qquad \ldots (3)$$

or

$$y = \mu + \delta + B_0 + B_{(1)} + B_{(2)} + \ldots + e \qquad \ldots (4)$$

where B is composed of contributions from variates B_0, $B_{(1)}$, $B_{(2)}$ ... and can account for a number of intermediate precision factors.

In practice, the objectives of a study and considerations of the sensitivity of the measurement method will govern the extent to which this model is used. In many cases, abbreviated forms will suffice.

6.4 Terms B_0, $B_{(1)}$, $B_{(2)}$, etc.

6.4.1 Under repeatability conditions, these terms all remain constant and add to the bias of the test results. Under intermediate precision conditions, B_0 is the fixed effect of the factor(s) that remained the same (state 1 of table 1), while $B_{(1)}$, $B_{(2)}$, etc. are the random effects of the factor(s) which vary (state 2 of table 1). These no longer contribute to the bias, but increase the intermediate precision standard deviation so that it becomes larger than the repeatability standard deviation.

6.4.2 The effects due to differences between operators include personal habits in operating measurement methods (e.g. in reading graduations on scales, etc.). Some of these differences should be removable by standardization of the measurement method, particularly in having a clear and accurate description of the techniques involved. Even though there is a bias in the test results obtained by an individual operator, that bias is not always constant (e.g. the magnitude of the bias will change according to his/her mental and/or physical conditions on that day) and the bias cannot be corrected or calibrated exactly. The magnitude of such a bias should be reduced by use of a clear operation manual and training. Under such circumstances, the effect of changing operators can be considered to be of a random nature.

6.4.3 The effects due to differences between equipment include the effects due to different places of installation, particularly in fluctuations of the indicator, etc. Some of the effects due to differences between equipment can be corrected by exact calibration. Differences due to systematic causes between equipment should be corrected by calibration, and such a procedure should be included in the standard method. For example, a change in the batch of a reagent could be treated that way. An accepted reference value is needed for this, for which ISO Guide 33 and ISO Guide 35 shall be consulted. The remaining effect due to equipment which has been calibrated using a reference material is considered a random effect.