INTERNATIONAL STANDARD

ISO 5725-4

First edition
1994-12-15

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

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

Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Partie 4: Méthodes de base pour la détermination de la justesse d'une méthode de mesure normalisée

Reference number
ISO 5725-4:1994(E)

Contents

Annexes

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 5725-4 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 and reproducibility conditions).

Annex A forms an integral part of this part of ISO 5725. Annexes B, C and D 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 arithmetic mean 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 has not been repeated in this part of ISO 5725. ISO 5725-1 should be read in conjunction with all other parts of ISO 5725, including this part, because it gives the underlying definitions and general principles.

0.3 The "trueness" of a measurement method is of interest when it is possible to conceive of a true value for the property being measured. Although, for some measurement methods, the true value cannot be known exactly, it may be possible to have an accepted reference value for the property being measured; for example, if suitable reference materials are available, or if the accepted reference value can be established by reference to another measurement method or by preparation of a known sample. The trueness of the measurement method can be investigated by comparing the accepted reference value with the level of the results given by the measurement method. Trueness is normally expressed in terms of bias. Bias can arise, for example, in chemical analysis if the measurement method fails to extract all of an element, or if the presence of one element interferes with the determination of another.

0.4 Two measures of trueness may be of interest and both are considered in this part of ISO 5725.

a) Bias of the measurement method: where there is a possibility that the measurement method may give rise to a bias, which persists wherever and whenever the measurement is done, then it is of interest to investigate the "bias of the measurement method" (as defined in ISO 5725-1). This requires an experiment involving many laboratories, very much as described in ISO 5725-2.

b) Laboratory bias: measurements within a single laboratory can reveal the "laboratory bias" (as defined in ISO 5725-1). If it is proposed to undertake an experiment to estimate laboratory bias, then it should be realized that the estimate will be valid only at the time of the experiment. Further regular testing is required to show that the laboratory bias does not vary; the method described in ISO 5725-6 may be used for this.

This page intentionally left blank

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

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

1 Scope

1.1 This part of ISO 5725 provides basic methods for estimating the bias of a measurement method and the laboratory bias when a measurement method is applied.

1.2 It 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 In order that the measurements are made in the same way, it is important that the measurement method has been standardized. All measurements are to be carried out according to that standard method.

1.4 Bias values give quantitative estimates of the ability of a measurement method to give the correct (true) result. When a value for the bias of a measurement method is quoted, together with a test result obtained by that method, there is an implication that the same characteristic is being measured in exactly the same way.

1.5 This part of ISO 5725 can be applied only if the accepted reference value can be established as a conventional true value, for example by measurement standards or suitable reference materials or by referring to a reference measurement method or by preparation of a known sample.

Reference materials could be either

a) certified reference materials;

b) materials manufactured for the purpose of the experiment with known properties; or

c) materials whose properties have been established by measurements using an alternative measurement method whose bias is known to be negligible.

1.6 This part of ISO 5725 considers only those cases where it is sufficient to estimate bias on one level at a time. It is not applicable if the bias in the measurement of one property is affected by the level of a second property (i.e. it does not consider interferences). Comparison of the trueness of two measurement methods is considered in ISO 5725-6.

NOTE 1 In this part of ISO 5725, bias is considered only at one level at a time. Therefore the index j for the level has been omitted throughout.

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.*

3 Definitions

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

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

4 Determination of the bias of a standard measurement method by an interlaboratory experiment

4.1 The statistical model

In the basic model described in subclause 5.1 of ISO 5725-1:1994, the general mean m may be replaced by

$$m = \mu + \delta \qquad \ldots (1)$$

where

μ is the accepted reference value of the property being measured;

δ is the bias of the measurement method.

The model becomes

$$y = \mu + \delta + B + e \qquad \ldots (2)$$

Equation (2) is used when δ is of interest. Here B is the laboratory component of bias, i.e. the component in a test result representing the between-laboratory variation.

The laboratory bias, Δ, is given by

$$\Delta = \delta + B \qquad \ldots (3)$$

so the model may be written

$$y = \mu + \Delta + e \qquad \ldots (4)$$

Equation (4) is used when Δ is of interest.

4.2 Reference material requirements

If reference materials are used, the requirements given in 4.2.1 and 4.2.2 shall be satisfied. Reference materials shall be homogeneous.

4.2.1 Choice of reference materials

4.2.1.1 The reference material shall have known properties at the level appropriate to the level at which the standard measurement method is intended to be applied, e.g. concentration. In some cases it will be important to include, in the assessment experiment, a series of reference materials, each corresponding to a different level of the property, as the bias of the standard measurement method may be different at different levels. The reference material should have a matrix as close as possible to the matrix of the material to be subjected to the standard measurement method, e.g. carbon in coal or carbon in steel.

4.2.1.2 The quantity of the reference material shall be sufficient for the entire experimental programme, including some in reserve if this is considered necessary.

4.2.1.3 Wherever possible, the reference material should have stable properties throughout the experiment. There are three cases, as follows.

a) The properties are stable: no precautions are necessary.

b) The certified value of the property may be influenced by storage conditions: the container should be stored, both before and after its opening, in the way described on the certificate.

c) The properties change at a known rate: there is a certificate supplied with the reference value to define the properties at specific times.

4.2.1.4 The possible difference between the certified value and the true value expressed by the uncertainty of the reference material (see ISO Guide 35) is not taken into account in the methods given here.

4.2.2 Check and distribution of the reference material

Where sub-division of the unit of the reference material occurs prior to distribution, it shall be performed with care to avoid the introduction of any additional error. Relevant International Standards on sample division should be consulted. The units should be selected on a random basis for distribution. If the measurement process is non-destructive, it is possible to give all the laboratories in the interlaboratory experiment the same unit of reference material, but this will extend the time-frame of the experiment.

4.3 Experimental design considerations when estimating the bias of a measurement method

4.3.1 The objective of the experiment is to estimate the magnitude of the bias of the measurement method and to determine if it is statistically significant. If the bias is found to be statistically insignificant, then the objective is to determine the magnitude of the maximum bias that would, with a certain probability, remain undetected by the results of the experiment.

4.3.2 The layout of the experiment is almost the same as that for a precision experiment, as described in subclause 4.1 of ISO 5725-2:1994. The differences are

a) there is an additional requirement to use an accepted reference value, and

b) the number of participating laboratories and the number of test results shall also satisfy the requirements given in 4.5.

4.4 Cross-references to ISO 5725-1 and ISO 5725-2

Clause 6 of ISO 5725-1:1994 and clauses 5 and 6 of ISO 5725-2:1994 apply. When reading parts 1 and 2 in this context, "trueness" should be inserted in place of "precision" or "repeatability and reproducibility" as appropriate.

4.5 Required number of laboratories

The number of laboratories and the number of test results required at each level are interdependent. The number of laboratories to be used is discussed in subclause 6.3 of ISO 5725-1:1994. A guide to deciding how many is given below.

In order for the results of an experiment to be able to detect with a high probability (see annex C) a predetermined magnitude of bias, the minimum number of laboratories, p, and test results, n, shall satisfy the following equation:

$$A\sigma_R \leqslant \frac{\delta_\mathrm{m}}{1{,}84} \qquad \ldots (5)$$

where

δ_m is the predetermined magnitude of bias that the experimenter wishes to detect from the results of the experiment;

σ_R is the reproducibility standard deviation of the measurement method.

A is a function of p and n and is given by

$$A = 1{,}96\sqrt{\frac{n\left(\gamma^2 - 1\right) + 1}{\gamma^2 pn}} \qquad \ldots (6)$$

where

$$\gamma = \sigma_R/\sigma_r \qquad \ldots (7)$$

Values of A are given in table 1.

Ideally, the choice of the combination of the number of laboratories and the number of replicate test results per laboratory should satisfy the requirement described by equation (5), with the δ_m value predetermined by the experimenter. However, for practical reasons, the choice of the number of laboratories is usually a compromise between the availability of resources and the desire to reduce the value of δ_m to a satisfactory level. If the reproducibility of the measurement method is poor, then it will not be practical to achieve a high degree of certainty in the estimate of the bias. When σ_R is larger than σ_r (i.e. γ is larger than 1) as is often the case, little is to be gained by obtaining more than $n = 2$ test results per laboratory per level.

Table 1 — Values showing the uncertainty in the estimate of the bias of the measurement method

p	$\gamma = 1$			$\gamma = 2$			$\gamma = 5$		
	$n = 2$	$n = 3$	$n = 4$	$n = 2$	$n = 3$	$n = 4$	$n = 2$	$n = 3$	$n = 4$
5	0,62	0,51	0,44	0,82	0,80	0,79	0,87	0,86	0,86
10	0,44	0,36	0,31	0,58	0,57	0,56	0,61	0,61	0,61
15	0,36	0,29	0,25	0,47	0,46	0,46	0,50	0,50	0,50
20	0,31	0,25	0,22	0,41	0,40	0,40	0,43	0,43	0,43
25	0,28	0,23	0,20	0,37	0,36	0,35	0,39	0,39	0,39
30	0,25	0,21	0,18	0,33	0,33	0,32	0,35	0,35	0,35
35	0,23	0,19	0,17	0,31	0,30	0,30	0,33	0,33	0,33
40	0,22	0,18	0,15	0,29	0,28	0,28	0,31	0,31	0,31

4.6 Statistical evaluation

The test results shall be treated as described in ISO 5725-2. In particular, if outlying values are detected, all necessary steps shall be taken to investigate the reasons why they have been obtained, including re-appraisal of the suitability of the accepted reference value.

4.7 Interpretation of the results of the statistical evaluation

4.7.1 Check of precision

The precision of the measurement method is expressed in terms of s_r (estimate of the repeatability standard deviation) and s_R (estimate of the reproducibility standard deviation). Equations (8) to (10) assume an equal number (n) of test results in each laboratory. If this is not true, the respective equations given in ISO 5725-2 should be used to calculate s_r and s_R.

4.7.1.1 The estimate s_r^2 of the repeatability variance for p participating laboratories is calculated as

$$s_r^2 = \frac{1}{p} \sum_{i=1}^{p} s_i^2 \qquad \dots (8)$$

$$s_i^2 = \frac{1}{n-1} \sum_{k=1}^{n} (y_{ik} - \bar{y}_i)^2 \qquad \dots (9)$$

$$\bar{y}_i = \frac{1}{n} \sum_{k=1}^{n} y_{ik} \qquad \dots (10)$$

where s_i^2 and $\bar{y}_i$ are respectively the variance and the average of n test results y_{ik} obtained in laboratory i.

Cochran's test, as described in ISO 5725-2, shall be applied to the variances s_i^2 to verify that no significant differences exist between the within-laboratory variances. Mandel's h and k plots, as described in ISO 5725-2, should also be drawn for a more thorough investigation of potential outliers.

If the repeatability standard deviation of the standard measurement method has not been previously determined in accordance with ISO 5725-2, s_r will be considered to be the best estimate of it. If the repeatability standard deviation of the standard test method, σ_r, has been determined in accordance with ISO 5725-2, s_r^2 can be assessed by computing the ratio

$$C = s_r^2 / \sigma_r^2 \qquad \dots (11)$$

The test statistic C is compared with the critical value

$$C_{\text{crit}} = \chi^2_{(1-\alpha)}(\nu)/\nu$$

where $\chi^2_{(1-\alpha)}(\nu)$ is the $(1-\alpha)$-quantile of the χ^2 distribution with $\nu\ [= p(n-1)]$ degrees of freedom. Unless otherwise stated, α is assumed to be 0,05.

a) If $C \leqslant C_{\text{crit}}$: s_r^2 is not significantly larger than σ_r^2.

b) If $C > C_{\text{crit}}$: s_r^2 is significantly larger than σ_r^2.

In the former case, the repeatability standard deviation, σ_r, will be used for the assessment of the bias of the measurement method. In the latter case, it is necessary to investigate the causes of the discrepancy and possibly to repeat the experiment prior to proceeding further.

4.7.1.2 The estimate, s_R^2, of the reproducibility variance for the p participating laboratories, is calculated as

$$s_R^2 = \frac{1}{p-1} \sum_{i=1}^{p} (\bar{y}_i - \bar{\bar{y}})^2 + \left(1 - \frac{1}{n}\right) s_r^2 \qquad \dots (12)$$

with

$$\bar{\bar{y}} = \frac{1}{p} \sum_{i=1}^{p} \bar{y}_i \qquad \ldots (13)$$

If the reproducibility standard deviation of the standard measurement method has not previously been determined in accordance with ISO 5725-2, s_R will be considered the best estimate of it. If the reproducibility standard deviation, σ_R, and the repeatability standard deviation, σ_r, of the standard measurement method have been determined in accordance with ISO 5725-2, s_R can be assessed indirectly by computing the ratio

$$C' = \frac{s_R^2 - (1 - 1/n)s_r^2}{\sigma_R^2 - (1 - 1/n)\sigma_r^2} \qquad \ldots (14)$$

The test statistic C' is compared with the critical value

$$C'_{\text{crit}} = \chi^2_{(1-\alpha)}(\nu)/\nu$$

where $\chi^2_{(1-\alpha)}(\nu)$ is the $(1-\alpha)$-quantile of the χ^2 distribution with ν $(= p - 1)$ degrees of freedom. Unless otherwise stated, α is assumed to be 0,05.

a) If $C' \leqslant C'_{\text{crit}}$: $s_R^2 - (1 - 1/n)s_r^2$ is not significantly larger than $\sigma_R^2 - (1 - 1/n)\sigma_r^2$.

b) If $C' > C'_{\text{crit}}$: $s_R^2 - (1 - 1/n)s_r^2$ is significantly larger than $\sigma_R^2 - (1 - 1/n)\sigma_r^2$.

In the former case, the repeatability standard deviation, σ_r, and the reproducibility standard deviation, σ_R, will be used for the assessment of the trueness of the measurement method. In the latter case, a careful examination of the working conditions of each laboratory shall be carried out before the assessment of the bias of the standard measurement method is undertaken. It may appear that some laboratories did not use the required equipment or did not work according to the specified conditions. In chemical analysis, problems may arise from, for example, insufficient control of temperature, moisture, presence of contaminants, etc. As a result the experiment may have to be repeated to yield the expected precision values.

4.7.2 Estimation of the bias of the standard measurement method

The estimate of the bias from the assessing laboratories is given by

$$\hat{\delta} = \bar{\bar{y}} - \mu \qquad \ldots (15)$$

where $\hat{\delta}$ may be positive or negative.

If the absolute value of the estimated bias is smaller than or equal to half the width of the uncertainty interval, as defined in ISO Guide 35, there is no evidence of a bias.

The variation of the estimate of the bias of the measurement method is due to the variation in the results of the measurement process and is expressed by its standard deviation computed as

$$\sigma_{\hat{\delta}} = \sqrt{\frac{\sigma_R^2 - (1 - 1/n)\sigma_r^2}{p}} \qquad \ldots (16)$$

in the case of known precision values, or

$$s_{\hat{\delta}} = \sqrt{\frac{s_R^2 - (1 - 1/n)s_r^2}{p}} \qquad \ldots (17)$$

in the case of unknown precision values.

An approximate 95 % confidence interval for the bias of the measurement method can be computed as

$$\hat{\delta} - A\sigma_R \leqslant \delta \leqslant \hat{\delta} + A\sigma_R \qquad \ldots (18)$$

where A is as given in equation (6). If σ_R is unknown, its estimate s_R has to be used instead, and A has to be computed with $\gamma = s_R/s_r$.

If this confidence interval covers the value zero, the bias of the measurement method is insignificant at the significance level $\alpha = 5$ %; otherwise it is significant.

5 Determination of the laboratory bias of one laboratory using a standard measurement method

As described below, experiments in one laboratory are used to estimate laboratory bias, provided that an interlaboratory precision experiment, in accordance with ISO 5725-2, has established the repeatability standard deviation of the method.

5.1 Carrying out the experiment

The experiment shall conform strictly to the standard method and measurements shall be carried out under repeatability conditions. Prior to conducting the assessment of trueness, a check of the precision of the standard measurement method as applied by the laboratory shall be performed. This implies comparison between the within-laboratory standard deviation and the stated repeatability standard deviation of the standard measurement method.

The layout of the experiment consists of the measurements required of one laboratory in a precision experiment as described in ISO 5725-2. Apart from the restriction to a single laboratory, the only substantial difference is the additional requirement to use an accepted reference value.

When attempting to measure the bias of a laboratory, it may not be worth putting a great deal of effort into such an experiment: the effort could perhaps be better expended by making checks at intervals as described in ISO 5725-6. If the repeatability of the measurement method is poor, then it will not be practical to achieve a high degree of certainty in the estimate of the bias of the laboratory.

5.2 Cross-references to ISO 5725-1 and ISO 5725-2

When reading ISO 5725-1 and ISO 5725-2 in this context, "trueness" should be inserted in place of "precision" or "repeatability and reproducibility" as appropriate. In ISO 5725-2, the number of laboratories will be $p = 1$, and it may be convenient for one person to combine the roles of "executive" and "supervisor".

5.3 Number of test results

The uncertainty in the estimate of the laboratory bias depends on the repeatability of the measurement method and on the number of test results obtained.

In order for the results of an experiment to be able to detect with a high probability (see annex C) a predetermined magnitude of bias, the number of test results, n, shall satisfy the following equation:

$$A_W \sigma_r \leqslant \frac{\Delta_m}{1,84} \qquad \ldots (19)$$

where

Δ_m is the predetermined magnitude of laboratory bias that the experimenter wishes to detect from the results of the experiment;

σ_r is the repeatability standard deviation of the measurement method and

$$A_W = \frac{1,96}{\sqrt{n}} \qquad \ldots (20)$$

5.4 Choice of reference materials

If a reference material is used, the requirements described in 4.2.1 also apply here.

5.5 Statistical analysis

5.5.1 Check of the within-laboratory standard deviation

Compute the average, $\bar{y}_W$, of the n test results and s_W, the estimate of the within-laboratory standard deviation σ_W, as follows:

$$\bar{y}_W = \frac{1}{n} \sum_{k=1}^{n} y_k \qquad \ldots (21)$$

$$s_W = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (y_k - \bar{y}_W)^2} \qquad \ldots (22)$$

The test results shall be scrutinized for outliers using Grubbs' test as described in subclause 7.3.4 of ISO 5725-2:1994.

If the repeatability standard deviation, σ_r, of the standard measurement method is known, the estimate s_W can be assessed by the following procedure.

Compute the ratio

$$C'' = (s_W/\sigma_r)^2 \qquad \ldots (23)$$

and compare the value C'' with the critical value

$$C''_{crit} = \chi^2_{(1-\alpha)}(\nu)/\nu$$

where $\chi^2_{(1-\alpha)}(\nu)$ is the $(1-\alpha)$-quantile of the χ^2 distribution with ν $[= n - 1]$ degrees of freedom. Unless otherwise stated, α is assumed to be 0,05.

a) If $C'' \leqslant C''_{crit}$: s_W is not significantly larger than σ_r.

b) If $C'' > C''_{crit}$: s_W is significantly larger than σ_r.

In the former case, the repeatability standard deviation of the measurement method, σ_r, will be used for the assessment of the laboratory bias.

In the latter case, consideration should be given to repeating the experiment with verification at all steps that the standard measurement method is properly implemented.

5.5.2 Estimation of the laboratory bias

The estimate, $\hat{\Delta}$, of the laboratory bias Δ is given by

$$\hat{\Delta} = \bar{y}_W - \mu \qquad \ldots (24)$$