SLOVENSKI STANDARD
SIST ISO 5725-6:2003

01-junij-2003

Točnost (pravilnost in natančnost) merilnih metod in rezultatov – 6. del : Uporaba vrednosti za točnost v praksi

Accuracy (trueness and precision) of measurement methods and results -- Part 6: Use in practice of accuracy values

Exactitude (justesse et fidélité) des résultats et méthodes de mesure -- Partie 6: Utilisation dans la pratique des valeurs d'exactitude

Ta slovenski standard je istoveten z: **ISO 5725-6:1994**

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 5725-6:2003 **en**

INTERNATIONAL STANDARD

ISO 5725-6

First edition
1994-12-15

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

Part 6:

Use in practice of accuracy values

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

Partie 6: Utilisation dans la pratique des valeurs d'exactitude

Reference number
ISO 5725-6:1994(E)

Contents

Page

1 Scope **1**

2 Normative references **1**

3 Definitions **2**

4 Determination of limits **2**

4.1 Repeatability and reproducibility limits **2**

4.2 Comparisons based on more than two values **3**

5 Methods for checking the acceptability of test results and determining the final quoted result **4**

5.1 General **4**

5.2 Methods for checking the acceptability of test results obtained under repeatability conditions **4**

5.3 Methods for checking the acceptability of test results obtained under both repeatability and reproducibility conditions **11**

6 Method for checking the stability of test results within a laboratory **13**

6.1 Background **13**

6.2 Methods for checking stability **13**

7 Use of repeatability and reproducibility standard deviations in assessing laboratories **25**

7.1 Assessment method **25**

7.2 Evaluation of the use of a measurement method by a laboratory not previously assessed **26**

7.3 Continued assessment of previously approved laboratories **29**

8 Comparison of alternative measurement methods **32**

8.1 Origin of alternative measurement methods **32**

8.2 Purpose of comparing measurement methods **32**

8.3 Method B is a candidate for an alternative standard method ("Standardization experiment" not defined) **33**

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

8.4 Accuracy experiment **33**

8.5 Method B is a candidate for a routine method **38**

Annex

A Symbols and abbreviations used in ISO 5725 ..... **40**

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 5725-6 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.

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 The need to consider "precision" arises because tests performed on presumably identical materials in presumably identical circumstances do not, in general, yield identical results. This is attributed to unavoidable random errors inherent in every measurement procedure; the factors that influence the outcome of a measurement cannot all be completely controlled. In the practical interpretation of measurement data, this variability has to be taken into account. For instance, the difference between a test result and some specified value may be within the scope of unavoidable random errors, in which case a real deviation from such a specified value has not been established. Similarly, comparing test results from two batches of material will not indicate a fundamental quality difference if the difference between them can be attributed to the inherent variation in the measurement procedure.

0.3 Parts 1 to 5 of ISO 5725 dicuss the background to, and given methods for, the assessment of the precision (in terms of the repeatability standard deviation and the reproducibility standard deviation) and the trueness (in terms of the various components of bias) of measurements produced by a standard measurement method. Such assessment would, however, be pointless if there were no practical uses to which the results could be put.

0.4 Given that the accuracy of a measurement method has been established, this part of ISO 5725 applies that knowledge in practical situations in such a way as to facilitate commercial transactions and to monitor and improve the operational performance of laboratories.

This page intentionally left blank

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

Part 6:
Use in practice of accuracy values

1 Scope

1.1 The purpose of this part of ISO 5725 is to give some indications of the way in which accuracy data can be used in various practical situations by:

a) giving a standard method of calculating the repeatability limit, the reproducibility limit and other limits to be used in examining the test results obtained by a standard measurement method;

b) providing a way of checking the acceptability of test results obtained under repeatability or reproducibility conditions;

c) describing how to assess the stability of results within a laboratory over a period of time, and thus providing a method of "quality control" of the operations within that laboratory;

d) describing how to assess whether a given laboratory is able to use a given standard measurement method in a satisfactory way;

e) describing how to compare alternative measurement methods.

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

1.3 It is assumed that the estimates of trueness and precision for the method have been obtained in accordance with parts 1 to 5 of ISO 5725.

1.4 Any additional information regarding the field of application will be given at the beginning of each particular application.

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

ISO 5725-4:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 4: Basic methods for the determination of the trueness of a standard measurement method.*

ISO 8258:1991, *Shewhart control charts.*

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

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

ISO/IEC Guide 25:1990, *General requirements for the competence of calibration and testing laboratories.*

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 Determination of limits

4.1 Repeatability and reproducibility limits

4.1.1 In ISO 5725-2, attention has been focussed on estimating the standard deviations associated with operations under repeatability or reproducibility conditions. However, normal laboratory practice requires examination of the difference(s) observed between two (or more) test results, and for this purpose some measure akin to a critical difference is required, rather than a standard deviation.

4.1.2 When a quantity is based on sums or differences of n independent estimates each having a standard deviation σ, then that resultant quantity will have a standard deviation $\sigma\sqrt{n}$. The reproducibility limit (R) or repeatability limit (r) are for differences between two test results, so the associated standard deviation is $\sigma\sqrt{2}$. In normal statistical practice, for examining the difference between these two values the critical difference used is f times this standard deviation, i.e. $f\sigma\sqrt{2}$. The value of f (the critical range factor) depends on the probability level to be associated with the critical difference and on the shape of the underlying distribution. For the reproducibility and repeatability limits, the probability level is specified as 95 %, and throughout the analysis in ISO 5725 the assumption is made that the underlying distribution is approximately normal. For a normal distribution at 95 % probability level, f is 1,96 and $f\sqrt{2}$ then is 2,77. As the purpose of this part of ISO 5725 is to give some simple "rule of thumb" to be applied by non-statisticians when examining the results of tests, it seems reasonable to use a rounded value of 2,8 instead of $f\sqrt{2}$.

4.1.3 As has been stated, the process of estimating precision leads to estimates of the true standard deviations while the true standard deviations remain unknown. Therefore in statistical practice they should be denoted by s rather than σ. However, if the procedures given in ISO 5725-1 and ISO 5725-2 are followed, these estimates will be based on an appreciable number of test results, and will give the best information we are likely to have of the true values of the standard deviations. In other applications that follow, for estimates of these standard deviations based on more limited data, the symbol s (estimate of a standard deviation) is used. Therefore it seems best to use the symbol σ to denote the values obtained from a full precision experiment, and treat these as true standard deviations with which other estimates (s) will be compared.

4.1.4 In view of 4.1.1 to 4.1.3, when examining two single test results obtained under repeatability or reproducibility conditions, the comparison shall be made with the repeatability limit

$$r = 2{,}8\sigma_r$$

or the reproducibility limit

$$R = 2{,}8\sigma_R$$

4.2 Comparisons based on more than two values

4.2.1 Two groups of measurements in one laboratory

If, in one laboratory under repeatability conditions, two groups of measurements are performed with the first group of n_1 test results giving an arithmetic mean of $\bar{y}_1$ and the second group of n_2 test results giving an arithmetic mean of $\bar{y}_2$, then the standard deviation of $(\bar{y}_1 - \bar{y}_2)$ is

$$\sigma = \sqrt{\sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

and the critical difference for $|\bar{y}_1 - \bar{y}_2|$ is

$$\text{CD} = 2{,}8\sigma_r \sqrt{\frac{1}{2n_1} + \frac{1}{2n_2}}$$

at the 95 % probability level.

NOTE 1 If n_1 and n_2 are both unity, this reduces to $r = 2{,}8\sigma_r$, as expected.

4.2.2 Two groups of measurements in two laboratories

If the first laboratory obtains n_1 test results giving an arithmetic mean of $\bar{y}_1$ while the second laboratory obtains n_2 test results giving an arithmetic mean of $\bar{y}_2$, in each case under repeatability conditions, then the standard deviation of $(\bar{y}_1 - \bar{y}_2)$ is

$$\sigma = \sqrt{\sigma_L^2 + \frac{1}{n_1}\sigma_r^2 + \sigma_L^2 + \frac{1}{n_2}\sigma_r^2}$$

$$= \sqrt{2\sigma_L^2 + \sigma_r^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$= \sqrt{2\left(\sigma_L^2 + \sigma_r^2\right) - 2\sigma_r^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)}$$

and the critical difference for $|\bar{y}_1 - \bar{y}_2|$ is

$$\text{CD} = \sqrt{(2{,}8\sigma_R)^2 - (2{,}8\sigma_r)^2 \left(1 - \frac{1}{2n_1} - \frac{1}{2n_2} \right)}$$

at the 95 % probability level.

NOTE 2 If n_1 and n_2 are both unity, this reduces to $R = 2{,}8\sigma_R$, as expected.

4.2.3 Comparison with a reference value for one laboratory

If n test results are obtained under repeatability conditions within one laboratory which give an arithmetic mean of $\bar{y}$, then the comparison with a given reference value μ_0 shall be made, in the absence of specific knowledge of the laboratory component of bias, using a standard deviation for $(\bar{y} - \mu_0)$ of

$$\sigma = \sqrt{\sigma_L^2 + \frac{1}{n}\sigma_r^2}$$

$$= \frac{1}{\sqrt{2}} \sqrt{2\left(\sigma_L^2 + \sigma_r^2\right) - 2\sigma_r^2 \left(1 - \frac{1}{n} \right)}$$

$$= \frac{1}{\sqrt{2}} \sqrt{2\left(\sigma_L^2 + \sigma_r^2\right) - 2\sigma_r^2 \left(\frac{n-1}{n} \right)}$$

and the critical difference for $|\bar{y} - \mu_0|$ is

$$\text{CD} = \frac{1}{\sqrt{2}} \sqrt{(2{,}8\sigma_R)^2 - (2{,}8\sigma_r)^2 \left(\frac{n-1}{n} \right)}$$

4.2.4 Comparison with a reference value for more than one laboratory

If p laboratories have obtained n_i test results giving arithmetic means of $\bar{y}_i$ (in each case under repeatability conditions) and the grand mean $\bar{\bar{y}}$ is computed by

$$\bar{\bar{y}} = \frac{1}{p} \sum \bar{y}_i$$

and this grand mean is to be compared with a reference value μ_0, then the standard deviation for $(\bar{\bar{y}} - \mu_0)$ is

$$\sigma = \sqrt{\frac{1}{p}\sigma_L^2 + \frac{1}{p^2}\sigma_r^2 \sum \frac{1}{n_i}}$$

$$= \frac{1}{\sqrt{2p}} \sqrt{2\left(\sigma_L^2 + \sigma_r^2\right) - 2\sigma_r^2 + \frac{2\sigma_r^2}{p} \sum \frac{1}{n_i}}$$

$$= \frac{1}{\sqrt{2p}} \sqrt{2\left(\sigma_L^2 + \sigma_r^2\right) - 2\sigma_r^2\left(1 - \frac{1}{p}\sum\frac{1}{n_i}\right)}$$

and the critical difference for $|\bar{\bar{y}} - \mu_0|$ is

$$\text{CD} = \frac{1}{\sqrt{2p}} \sqrt{(2{,}8\sigma_R)^2 - (2{,}8\sigma_r)^2\left(1 - \frac{1}{p}\sum\frac{1}{n_i}\right)}$$

at the 95 % probability level.

4.2.5 Quoting the results of a comparison

When the absolute difference exceeds the appropriate limit as given in the preceding clauses, then the difference shall be considered as suspect, and therefore all measurements that have given rise to this difference shall be considered as suspect and subject to further investigation.

5 Methods for checking the acceptability of test results and determining the final quoted result

5.1 General

5.1.1 The checking method described in this clause should be applied only to the case where the measurement was carried out according to a measurement method which has been standardized and whose standard deviations σ_r and σ_R are known. Therefore, when the range of N test results exceeds the appropriate limit as given in clause 4, it is considered that one, two or all of the N test results is or are aberrant. It is recommended that the cause of the aberrant result(s) should be investigated from the technical point of view. However, it may be necessary for commercial reasons to obtain some acceptable value, and in such cases the test results shall be treated according to the stipulations of this clause.

5.1.2 This clause has been prepared on the assumptions that the test results were obtained under repeatability and reproducibility conditions, and that the probability level to be used is 95 %. If intermediate conditions (see ISO 5725-3) were in force, then it is necessary to replace σ_r by the appropriate intermediate measure.

5.1.3 In some cases where the procedures described in 5.2 lead to the median being quoted as the final result, it might be better to abandon the data.

5.2 Methods for checking the acceptability of test results obtained under repeatability conditions

NOTE 3 In 5.2.2.1 and 5.2.2.2, reference made to measurements being expensive or inexpensive should be interpreted not only in financial terms but also whether the measurement is complex, troublesome or time-consuming.

5.2.1 Single test result

It is not common in commercial practice to obtain only one test result. When only one test result is obtained, it is not possible to make an immediate statistical test of the acceptability of that test result with respect to the given repeatability measure. If there is any suspicion that the test result may not be correct, a second test result should be obtained. Availability of two test results leads to the more common practice which is described below.

5.2.2 Two test results

The two test results should be obtained under repeatability conditions. The absolute difference between the two test results should then be compared with the repeatability limit $r = 2{,}8\sigma_r$.

5.2.2.1 Case where obtaining test results is inexpensive

If the absolute difference between the two test results does not exceed r, then both test results are considered acceptable, and the final quoted result should be quoted as the arithmetic mean of the two test results. If the absolute difference does exceed r, the laboratory should obtain two further test results.

If the range $(x_{max} - x_{min})$ of the four test results is equal to or less than the critical range at the 95 % probability level for $n = 4$, $\text{CR}_{0,95}(4)$, the arithmetic mean of the four test results should be reported as the final quoted result. Critical range factors, $f(n)$, for $n = 2$ to $n = 40$ and selected values from $n = 45$ to $n = 100$ are given in table 1 to be used to calculate the critical range according to the following equation:

$$\text{CR}_{0,95}(n) = f(n)\sigma_r$$

If the range of the four test results is greater than the critical range for $n = 4$, the median of the four test results should be reported as the final quoted result.

This procedure is summarized in the flowchart given in figure 1.

5.2.2.2 Case where obtaining test results is expensive

If the absolute difference between the two test results does not exceed r, then both test results are considered acceptable, and the final quoted result should be quoted as the arithmetic mean of the two test results. If the absolute difference does exceed r, the laboratory should obtain a further test result.

If the range $(x_{max} - x_{min})$ of the three test results is equal to or less than the critical range for $n = 3$, $CR_{0,95}(3)$, the arithmetic mean of the three test results should be reported as the final quoted result.

If the range of the three test results is greater than the critical range for $n = 3$, a decision on one of the following two cases shall be made.

a) **Case where it is impossible to obtain a fourth test result:**

The laboratory should use the median of the three test results as the final quoted result.

This procedure is summarized in the flowchart given in figure 2.

b) **Case where it is possible to obtain a fourth test result:**

The laboratory should obtain the fourth test result. If the range $(x_{max} - x_{min})$ of the four test results is equal to or less than the critical range for $n = 4$, $CR_{0,95}(4)$, the arithmetic mean of the four test results should be reported as the final quoted result. If the range of the four test results is greater than the critical range for $n = 4$, the laboratory should use the median of the four test results as the final quoted result.

This procedure is summarized in the flowchart given in figure 3.

Table 1 — Critical range factors, $f(n)$

n	$f(n)$	n	$f(n)$
2	2,8	25	5,2
3	3,3	26	5,2
4	3,6	27	5,2
5	3,9	28	5,3
6	4,0	29	5,3
7	4,2	30	5,3
8	4,3	31	5,3
9	4,4	32	5,3
10	4,5	33	5,4
11	4,6	34	5,4
12	4,6	35	5,4
13	4,7	36	5,4
14	4,7	37	5,4
15	4,8	38	5,5
16	4,8	39	5,5
17	4,9	40	5,5
18	4,9	45	5,6
19	5,0	50	5,6
20	5,0	60	5,8
21	5,0	70	5,9
22	5,1	80	5,9
23	5,1	90	6,0
24	5,1	100	6,1

NOTE — The critical range factor $f(n)$ is the 95 % quantile of the distribution of $(x_{max} - x_{min})/\sigma$ where x_{max} and x_{min} are the extreme values in a sample of size n from a normal distribution with standard deviation σ.