BS ISO 5725-2:2019

BSI Standards Publication

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

Documents
Document Sharing Hub

Full document details: BS ISO 5725-2:2019

National foreword

This British Standard is the UK implementation of ISO 5725-2:2019. It supersedes BS ISO 5725-2:1994, which is withdrawn.

The UK participation in its preparation was entrusted to Technical Committee SS/6/-/1, Precision of test methods - Panel 1: Revision of ISO 5725.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2019
Published by BSI Standards Limited 2019

ISBN 978 0 580 91453 9

ICS 03.120.30; 17.020

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 December 2019.

Amendments/corrigenda issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: BS ISO 5725-2:2019

INTERNATIONAL STANDARD

ISO 5725-2

Second edition
2019-12

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

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

Partie 2: Méthode de base pour la détermination de la répétabilité et de la reproductibilité d'une méthode de mesure normalisée

Reference number
ISO 5725-2:2019(E)

Documents
Document Sharing Hub

© ISO 2019

Full document details: BS ISO 5725-2:2019

COPYRIGHT PROTECTED DOCUMENT

© ISO 2019

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO 5725-2:2019

Contents

Page

Foreword ... **v**
Introduction ... **vi**
1 Scope ... **1**
2 Normative references ... **2**
3 Terms and definitions ... **2**
4 Symbols ... **2**
5 Estimates of the parameters in the basic model ... **4**
6 Requirements for a precision experiment ... **5**
6.1 Layout of the experiment ... 5
6.2 Recruitment of the laboratories ... 6
6.3 Preparation of the materials ... 6
7 Personnel involved in a precision experiment ... **7**
7.1 Panel ... 7
7.2 Statistical functions ... 8
7.3 Executive functions ... 8
7.4 Supervisors ... 9
7.5 Operators ... 10
8 Statistical analysis of a precision experiment ... **10**
8.1 Preliminary considerations ... 10
8.2 Tabulation of the results and notation used ... 11
8.2.1 Cells ... 11
8.2.2 Redundant data ... 11
8.2.3 Missing data ... 11
8.2.4 Outliers ... 11
8.2.5 Outlying laboratories ... 11
8.2.6 Erroneous data ... 11
8.2.7 Balanced uniform-level test results ... 11
8.2.8 Collation of data and intermediate values ... 12
8.2.9 Original test results ... 12
8.2.10 Cell means (Form B of Figure 2) ... 12
8.2.11 Measures of cell spread (Form C of Figure 2) ... 12
8.2.12 Corrected or rejected data ... 13
8.3 Scrutiny of results for consistency and outliers ... 13
8.3.1 Approaches for scrutiny of data ... 13
8.3.2 Graphical consistency technique ... 13
8.3.3 Numerical outlier technique ... 16
8.3.4 Cochran's test ... 16
8.3.5 Grubbs' tests ... 18
8.3.6 Repeated testing for outlying means or outlying data points ... 20
8.3.7 Alternative outlier inspection and test methods ... 20
8.4 Calculation of the general mean and variances ... 20
8.4.1 Method of analysis ... 20
8.4.2 Basic data ... 21
8.4.3 Non-empty cells ... 21
8.4.4 Calculation of the general mean, $\hat{m}$ ... 21
8.4.5 Calculation of variances ... 21
8.4.6 Alternative calculation methods for variances ... 22
8.4.7 Dependence of the variances upon m ... 23
8.5 Establishing a functional relationship between precision values, s, and the mean level, m ... 23
8.5.1 Choice of functional relationship ... 23

AVA Documents
Document Sharing Hub

© ISO 2019 – All rights reserved

Full document details: BS ISO 5725-2:2019

8.5.2 Fitting relationships I and II ... 24
8.5.3 Fitting relationship III ... 25
8.5.4 Fitting relationship IV ... 26
8.6 Statistical analysis as a step-by-step procedure ... 28
8.7 Report to the panel and decisions to be taken by the panel ... 30
8.7.1 Report by the statistical expert ... 30
8.7.2 Decisions to be taken by the panel ... 32
8.7.3 Full report ... 33
9 Statistical tables ... 33
Annex A (informative) **Number of laboratories required for an estimate of precision ... 38**
Annex B (informative) **Alternative calculations of variance components ... 41**
Annex C (informative) **Examples of the statistical analysis of precision experiments ... 44**
Annex D (informative) **Calculation of critical values and indicators ... 66**
Bibliography ... 69

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

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 documents 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).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

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 5725-2:1994), which has been technically revised. It also incorporates the Technical Corrigendum ISO 5725-2:1994/Cor 1:2002.

The main changes compared to the previous edition are as follows:

— permission is given to use alternative scrutiny and outlier detection tests provided that the performance is similar;

— permission is given to apply modern statistical methods available for calculations of the relevant precision and trueness characteristics;

— guidance on the number of laboratories required for a precision study has been included;

— information on the computation of critical values has been included.

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.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

Introduction

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.

General consideration of these quantities is given in ISO 5725-1 and so is not repeated in this document. 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.

This document is concerned solely with estimating the repeatability standard deviation and reproducibility standard deviation based on an interlaboratory design in which each laboratory conducts a number of independent measurements of the same sample under repeatability conditions. There are other designs (such as nested, factorial or split-level experiments) which can be used for the estimation of precision: these are not dealt with in this document but rather are the subject of other parts of ISO 5725. Nor does this document consider any other measures of precision intermediate between the two principal measures; those are the subject of ISO 5725-3.

In certain circumstances, the data obtained from an experiment carried out to estimate precision are used also to estimate trueness and can be used to evaluate measurement uncertainty. The estimation of trueness is not considered in this document; all aspects of the estimation of trueness are the subject of ISO 5725-4. The evaluation of measurement uncertainty, using inter-laboratory estimates of trueness and precision, is the subject of ISO 21748.

Annex C provides practical examples of estimating the precision of measurement methods by experiment. Worked examples are given to demonstrate balanced uniform sets of test results, although in one example a variable number of replicates per cell were reported (unbalanced design) and in another some data were missing. This is because an experiment designed to be balanced can turn out to be unbalanced. Stragglers and outliers are also considered.

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

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

1 Scope

1.1 This document

— amplifies the general principles for designing experiments for the numerical estimation of the precision of measurement methods by means of a collaborative interlaboratory experiment;

— provides a detailed practical description of the basic method for routine use in estimating the precision of measurement methods;

— provides guidance to all personnel concerned with designing, performing or analysing the results of the tests for estimating precision.

NOTE Modifications to this basic method for particular purposes are given in other parts of ISO 5725.

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 this single value can be the outcome of a calculation from a set of observations.

1.3 It assumes that in the design and performance of the precision experiment, all the principles as laid down in ISO 5725-1 are observed. The basic method uses the same number of test results in each laboratory, with each laboratory analysing the same levels of test sample; i.e. a balanced uniform-level experiment. The basic method applies to procedures that have been standardized and are in regular use in a number of laboratories.

1.4 The statistical model of ISO 5725-1:1994, Clause 5, is accepted as a suitable basis for the interpretation and analysis of the test results, the distribution of which is approximately normal.

1.5 The basic method, as described in this document, (usually) estimates the precision of a measurement method:

a) when it is required to determine the repeatability and reproducibility standard deviations as defined in ISO 5725-1;

b) when the materials to be used are homogeneous, or when the effects of heterogeneity can be included in the precision values; and

c) when the use of a balanced uniform-level layout is acceptable.

1.6 The same approach can be used to make a preliminary estimate of precision for measurement methods which have not reached standardization or are not in routine use.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

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: Probability and general statistical terms*

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

ISO 3534-3, *Statistics — Vocabulary and symbols — Part 3: Design of experiments*

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

3 Terms and definitions

For the purposes of this document, the definitions given in ISO 3534-1, ISO 3534-2, ISO 3534-3, and ISO 5725-1 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 http://www.electropedia.org/

4 Symbols

α	Probability associated with a critical value of a test statistic, also referred to as a level of significance
a	Intercept in the relationship $s = a + bm$
a_v	Intercept parameter in the relationship $s_j^2 = a_v^2 + (b_v m)^2$
A	Factor used to calculate the uncertainty of an estimate
b	Slope in the relationship $s = a + bm$
b_v	Slope parameter in the relationship $s_j^2 = a_v^2 + (b_v m)^2$
B	Laboratory component of bias under repeatability conditions
c	Intercept in the relationship $\lg s = c + d \lg m$
C, C', C''	Test statistics
C_{crit}, C'_{crit}, C''_{crit}	Critical values for statistical tests
d	Slope in the relationship $\lg s = c + d \lg m$
e	Component in a test result representing the random error occurring in every test result
G	Grubbs' test statistic
h	Mandel's between-laboratory consistency test statistic
k	Mandel's within-laboratory consistency test statistic

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

$L(\theta)$	Log-likelihood for variance components θ
m	General mean of the test property; level
$\hat{m}$	Estimate of the general mean of the test property
$\mathbf{M}$	Transformation matrix used in REML estimation
N	Number of iterations
n	Number of test results obtained in one laboratory at one level (i.e. per cell)
n_j	Total number of test results obtained at level j of the interlaboratory experiment
p	Number of laboratories participating in the interlaboratory experiment
P	Probability
q	Number of levels of the test property in the interlaboratory experiment
r	Repeatability limit
R	Reproducibility limit
s	Estimate of a standard deviation
$\hat{s}$	Predicted standard deviation
T	Total or sum of some expression
t	Number of test objects or groups
$\mathbf{V}(\boldsymbol{\theta})$	Covariance matrix used in REML estimation
W	Weighting factor used in calculating a weighted regression
w	Weighting factor used in calculating a weighted mean
x	Datum used for Grubbs' test
$\mathbf{X}$	Design matrix for REML estimations
y	Test result
$\bar{\bar{y}}$	Grand mean of test results
$\mathbf{Y}$	Vector of all observations at a level j
θ	Vector of variance components used in REML estimation
μ	True value or accepted reference value of a test property
σ	True value of a standard deviation

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

Subscripts

i	Identifier for a particular laboratory Index for summation (Annex A)
j	Identifier for a particular level Index for summation (Annex A)
k	Identifier for a particular test result in a laboratory *i* at level *j*
L	Between-laboratory (interlaboratory)
P	Probability
r	Repeatability
R	Reproducibility
REML	Estimate arising from a restricted maximum likelihood calculation
v	Terms used in calculation of a relationship between mean and combined variance (see 8.5.1.3, relationship III)
W	Within-laboratory (intralaboratory)
1, 2, 3, …	For test results, numbering in the order of obtaining them; for other cases (laboratories), as arbitrary identifiers
(1), (2), (3), …	For test results, (1), (2) … denote the 1st, 2nd … etc. order statistic, that is, the 1st, 2nd… etc. value numbered in the order of increasing magnitude

5 Estimates of the parameters in the basic model

5.1 The procedures given in this document are based on the statistical model given in Clause 5 of ISO 5725-1:1994 and elaborated upon in ISO 5725-1:1994, 1.2. In particular, these procedures are based on Formulae (2) to (6) of ISO 5725-1:1994, Clause 5.

The model is

$$y = m + B + e$$

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.

NOTE The laboratory component of bias, *B*, represents the deviation of a laboratory mean from the general average *m*.

5.2 ISO 5725-1:1994, Formulae (2) to (6), are expressed in terms of the true standard deviations of the populations considered. In practice, the exact values of these standard deviations are not known, and estimates of precision values must be made from a relatively small sample of all the possible laboratories, and within those laboratories from a small sample of all the possible test results.

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

5.3 In statistical practice, where the true value of a standard deviation, σ, is not known and is replaced by an estimate based upon a sample, then the symbol σ is replaced by s to denote that it is an estimate. This is done in each of ISO 5725-1:1994, Formulae (2) to (6), giving:

— s_L^2 is the estimate of the between-laboratory variance;

— s_W^2 is the estimate of the within-laboratory variance;

— s_r^2 is the arithmetic mean of s_W^2 and is the estimate of the repeatability variance; this arithmetic mean is taken over all those laboratories taking part in the accuracy experiment which remain after outliers have been excluded;

— s_R^2 is the estimate of the reproducibility variance:

$$s_R^2 = s_L^2 + s_r^2 \tag{1}$$

6 Requirements for a precision experiment

6.1 Layout of the experiment

6.1.1 In the layout used in the basic method, samples from q batches of materials, representing q different levels of the test, are sent to p laboratories which each obtain exactly n replicate test results under repeatability conditions at each of the q levels. This type of experiment is called a balanced uniform-level experiment.

6.1.2 The performance of these measurements shall be organized and instructions issued as follows.

a) Any preliminary checking of equipment shall be as specified in the standard method.

b) Each group of n measurements belonging to one level shall be carried out under repeatability conditions, i.e. within a short interval of time and by the same operator, and without any intermediate recalibration of the apparatus unless this is an integral part of performing a measurement.

c) It is essential that a group of n tests under repeatability conditions be performed independently as if they were n tests on different materials. As a rule, however, the operator knows that he/she is testing identical material, but the point should be stressed in the instructions that the whole purpose of the experiment is to determine what differences in results can occur in actual testing. If it is feared that, despite this warning, previous results can influence subsequent test results and thus the repeatability variance, it should be considered whether to use n separate samples at each of the q levels, coded in such a way that the operator does not know which are the replicates for a given level. However, such a procedure can cause problems in ensuring that repeatability conditions apply between replicates. This is only possible if the measurements are of such a nature that all the qn measurements can be performed within a short interval of time.

d) It is not essential that all the q groups of n measurements each be performed strictly within a short interval of time; different groups of measurements may be carried out on different days.

e) Measurements of all q levels shall be performed by one and the same operator and, in addition, the n measurements at a given level shall be performed using the same equipment throughout.

f) If in the course of the measurements an operator should become unavailable, another operator may complete the measurements, provided that the change does not occur within a group of n measurements at one level but only occurs between two of the q groups. Any such change shall be reported with the results.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

g) A time limit shall be given within which all measurements shall be completed. This can be necessary to limit the time allowed to elapse between the day the samples are received and the day the measurements are performed.

h) All samples shall be clearly labelled with the name of the experiment and a sample identification.

6.1.3 In 6.1.2 and elsewhere in this document, reference is made to the operator. For some measurements, there can in fact be a team of operators, each of whom performs some specific part of the procedure. In such a case, the team shall be regarded as "the operator" and any change in the team shall be regarded as providing a different "operator".

6.1.4 In commercial practice, the test results can be rounded rather crudely, but in a precision experiment test results shall be reported to at least one more digit than specified in the standard method. If the method does not specify the number of digits, the rounding shall not be coarser than half the repeatability standard deviation estimate. When precision depends on the level *m*, different degrees of rounding can be necessary for different levels.

6.2 Recruitment of the laboratories

6.2.1 The general principles regarding recruitment of the laboratories to participate in an interlaboratory experiment are given in ISO 5725-1. Guidance on the number of laboratories is given in Annex A. In enlisting the cooperation of the requisite number of laboratories, their responsibilities shall be clearly stated. An example of a suitable enlistment questionnaire is given in Figure 1.

6.2.2 For the purposes of this document, a "laboratory" is considered to be a combination of the operator, the equipment and the test site. One test site (or laboratory in the conventional sense) can thus produce several "laboratories" if it can provide several operators each with independent sets of equipment and situations in which to perform the work.

6.3 Preparation of the materials

6.3.1 A discussion of the points that need to be considered when selecting materials for use in a precision experiment is given in ISO 5725-1.

6.3.2 When deciding on the quantities of material to be provided, allowance shall be made for accidental spillage or errors in obtaining some test results which can necessitate using extra material. The amount of material prepared shall be sufficient to cover the experiment and allow an adequate stock in reserve.

6.3.3 It should be considered whether it is desirable for some laboratories to obtain some preliminary test results for familiarization with the measurement method before obtaining the official test result and, if so, whether additional material (not precision experiment samples) should be provided for this purpose.

6.3.4 When a material is to be homogenized, this shall be done in the manner most appropriate for that material. When the material to be tested is not homogeneous, it is important to prepare the samples in the manner specified in the method, preferably starting with one batch of commercial material for each level. In the case of unstable materials, special instructions on storage and treatment shall be specified.

NOTE ISO Guide 35 gives information on evaluating homogeneity and stability for reference materials.

6.3.5 For the samples at each level, *n* separate containers shall be used for each laboratory if there is any danger of the materials deteriorating once the container has been opened (e.g. by oxidation, by losing volatile components, or with hygroscopic material). In the case of unstable materials, special instructions on storage and treatment shall be specified. Precautions can be necessary to ensure that samples

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

remain identical up to the time the measurements are made. If the material to be measured consists of a mixture of powders of different relative density or of different grain size, some care is needed because segregation can result from shaking, for example during transport. When reaction with the atmosphere can be expected, the specimens may be sealed into ampoules, either evacuated or filled with an inert gas. For perishable materials such as food or blood samples, it can be necessary to send them in a deep-frozen state to the participating laboratories with detailed instructions for the procedure for thawing.

Questionnaire for interlaboratory study

Title of measurement method (copy attached) .....

1. Our laboratory is willing to participate in the precision experiment for this standard measurement method

   YES ☐ NO ☐ (tick appropriate box)

2. As a participant, we understand that:
   a) all essential apparatus, chemicals, and other requirements specified in the method must be available in our laboratory when the programme begins;
   b) specified "timing" requirements such as starting date, order of testing specimens and finishing date of the programme must be rigidly met;
   c) the method must be strictly adhered to;
   d) samples must be handled in accordance with instructions;
   e) a qualified operator must perform the measurements.

Having studied the method and having made a fair appraisal of our capabilities and facilities, we feel that we will be adequately prepared for cooperative testing of this method.

3. Comments

(Signed) .....

(Company or laboratory) .....

Figure 1 — Enlistment questionnaire for interlaboratory study

7 Personnel involved in a precision experiment

NOTE The methods of operation within different laboratories are not expected to be identical. Therefore, the contents of this clause are only intended as a guide to be modified as appropriate to cater for a particular situation.

7.1 Panel

7.1.1 The precision experiment should be overseen by a panel which should consist of experts familiar with the measurement method and its application.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

7.1.2 The tasks of the panel are:

a) to plan and coordinate the precision experiment;

b) to decide on the number of laboratories, levels and measurements to be made, and the number of significant digits to be required;

c) to appoint someone for the statistical functions (see 7.2);

d) to appoint someone for the executive functions (see 7.3);

e) to consider the instructions to be issued to the laboratory supervisors in addition to the standard measurement method;

f) to decide whether some operators can be allowed to carry out a few unofficial measurements in order to regain experience of the method after a long interval (such measurements shall never be carried out on the official collaborative samples);

g) to discuss the report of the statistical analysis on completion of the analysis of the test results;

h) to establish final values for the repeatability standard deviation and the reproducibility standard deviation;

i) to decide if further actions are required to improve the standard for the measurement method or with regard to laboratories whose test results have been rejected as outliers.

7.2 Statistical functions

At least one member of the panel should have experience in statistical design and analysis of experiments. His/her tasks are:

a) to contribute his/her specialized knowledge in designing the experiment;

b) to analyse the data;

c) to write a report for submission to the panel following the instructions contained in 8.7.

7.3 Executive functions

7.3.1 The actual organization of the experiment should be entrusted to a single laboratory. A member of the staff of that laboratory should take full responsibility; he/she is called the executive officer and is appointed by the panel.

7.3.2 The tasks of the executive officer are:

a) to enlist the cooperation of the requisite number of laboratories and to ensure that supervisors are appointed;

b) to organize and supervise the preparation of the materials and samples and the dispatch of the samples; for each level, an adequate quantity of material should be set aside as a reserve stock;

c) to draft instructions covering all the points in 6.1.2 a) to h), and circulate them to the supervisors early enough in advance for them to raise any comments or queries and to ensure that operators selected are those who normally carry out such measurements in routine operations;

d) to design suitable forms for the operator to use as a working record and for the supervisor to report the test results to the requisite number of decimal places (or significant digits, as required). Such forms can include the name of the operator, the dates on which samples were received and measured, the equipment used and any other relevant information;

e) to deal with any queries from laboratories regarding the performance of the measurements;

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

f) to see that an overall time schedule is maintained;

g) to collect the data forms and present them to the statistical expert.

NOTE The forms referred to in 7.3.2 d) can be electronic; for example a spreadsheet format suitably protected against unintended modification.

7.4 Supervisors

7.4.1 A staff member in each of the participating laboratories should be made responsible for organizing the actual performance of the measurements, in keeping with instructions received from the executive officer, and for reporting the test results.

7.4.2 The tasks of the supervisor are:

a) to ensure that the operators selected are those who normally carry out such measurements in routine operations;

b) to hand out the samples to the operator(s) in keeping with the instructions of the executive officer (and to provide material for familiarization experiments, if necessary);

c) to supervise the execution of the measurements (the supervisor shall not take part in performing the measurements);

d) to ensure that the operators carry out the required number of measurements;

e) to ensure adherence to the set timetable for performing the measurements;

f) to collect the test results recorded to the agreed number of decimal places (or significant digits), including any anomalies and difficulties experienced, and comments made by the operators.

7.4.3 The supervisor of each laboratory should write a full report which should contain the following information:

a) the test results, entered legibly by their originator on the forms provided, not transcribed or typed (computer or testing machine output may be acceptable as an alternative);

b) the original observed values or readings (if any) from which the test results were derived, entered legibly by the operator on the forms provided, not transcribed or typed;

c) comments by the operators on the standard for the measurement method;

d) information about irregularities or disturbances that can have occurred during the measurements, including any change of operator, together with a statement as to which measurements were performed by which operator, and the reasons for any missing results;

e) the date(s) on which the samples were received;

f) the date(s) on which each sample was measured;

g) information about the equipment used, if relevant;

h) any other relevant information.

NOTE The output and forms referred to in 7.4.3 a) and b) can be electronic; for example a spreadsheet format suitably protected against unintended modification.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

7.5 Operators

7.5.1 In each laboratory the measurements shall be carried out by one operator selected as being representative of those likely to perform the measurements in normal operations.

7.5.2 Because the object of the experiment is to determine the precision obtainable by the general population of operators working from the standard measurement method, in general the operators should not be given amplifications to the standard for the measurement method. However, it should be pointed out to the operators that the purpose of the exercise is to discover the extent to which results can vary in practice, so that there is less temptation for them to discard or rework results that they feel are inconsistent.

7.5.3 Although normally the operators should receive no supplementary amplifications to the standard measurement method, they should be encouraged to comment on the standard and, in particular, to state whether the instructions contained in it are sufficiently unambiguous and clear.

7.5.4 The tasks of the operators are:

a) to perform the measurements according to the standard measurement method;

b) to report any anomalies or difficulties experienced; it is better to report a mistake than to adjust the test results because one or two missing test results do not spoil the experiment and many indicate a deficiency in the standard;

c) to comment on the adequacy of the instructions in the standard; operators should report any occasion(s) when they are unable to follow their instructions as this can also indicate a deficiency in the standard.

8 Statistical analysis of a precision experiment

8.1 Preliminary considerations

8.1.1 The analysis of the data, which should be considered as a statistical problem to be solved by a statistical expert, involves three successive stages:

a) critical examination of the data in order to identify and treat outliers or other irregularities and to test the suitability of the model;

b) computation of preliminary values of precision and means for each level separately;

c) establishment of final values of precision and means, including the establishment of a relationship between precision and the level m when the analysis indicates that such a relationship can exist.

8.1.2 The analysis first computes, for each level separately, estimates of

— the repeatability variance, s_r^2

— the between-laboratory variance, s_L^2

— the reproducibility variance, s_R^2

— the mean, m.

NOTE The analysis includes a systematic application of statistical tests for outliers, a great variety of which are available from the literature and which can be used for the purposes of this document. For practical reasons, only a limited number of these tests, as explained in 8.3, have been incorporated in this document.

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

8.2 Tabulation of the results and notation used

8.2.1 Cells

Each combination of a laboratory and a level is called a cell of the precision experiment. In the ideal case, the results of an experiment with p laboratories and q levels consist of a table with pq cells, each containing n replicate test results that can all be used for computing the repeatability standard deviation and the reproducibility standard deviation. This ideal situation is not, however, always attained in practice. Departures occur owing to redundant data, missing data and outliers.

8.2.2 Redundant data

Sometimes a laboratory can carry out and report more than the n test results officially specified. In that case, the supervisor shall report why this was done and which are the correct test results. If the answer is that they are all equally valid, then a random selection may be made from those available test results to choose the planned number of test results for analysis.

NOTE The calculations of precision terms that are permitted in this document can accommodate differing numbers of test results in each cell (see 8.2.7). However, substantial variation in the number of reported observations can adversely affect the interpretation of outlier tests and Mandel's statistics and can reduce the reliability of the grand mean.

8.2.3 Missing data

In other cases, some of the test results can be missing, for example because of loss of a sample or a mistake in performing the measurement. The analysis recommended in this document is such that completely empty cells can simply be ignored, while partly empty cells can be taken into account by the standard computational procedure.

8.2.4 Outliers

These are entries among the original test results, or in the tables derived from them, that deviate so much from the comparable entries in the same table that they are considered irreconcilable with the other data. Experience has taught that outliers cannot always be avoided and they have to be taken into consideration in a similar way to the treatment of missing data.

8.2.5 Outlying laboratories

When several unexplained abnormal test results occur at different levels within the same laboratory, then that laboratory can be considered to be an outlier, having too high a within-laboratory variance and/or too large a systematic error in the level of its test results. It can then be reasonable to discard some or all of the data from such an outlying laboratory.

This document does not provide a statistical test by which suspected laboratories can be judged. The primary decision should be the responsibility of the statistical expert, but all rejected laboratories shall be reported to the panel for further action.

8.2.6 Erroneous data

Obviously erroneous data should be investigated and corrected or discarded.

NOTE Erroneous data can be obvious when, for example, results are reported in incorrect units, the values reported cannot reasonably relate to the test material under consideration, or other reported information indicates substantial unintended deviation from the standard method under study.

8.2.7 Balanced uniform-level test results

The ideal case is p laboratories called i (i = 1, 2..., p), each testing q levels called j (j = 1, 2, ..., q) with n replicates at each level (each ij combination), giving a total of pqn test results. Because of missing (8.2.3)

© ISO 2019 – All rights reserved

AVA Documents Document Sharing Hub
Full document details: BS ISO 5725-2:2019

or deviating (8.2.4) test results, or outlying laboratories (8.2.5) or erroneous data (8.2.6), this ideal situation is not always attained. Under these conditions the notations given in 8.2.9 to 8.2.11 and the procedures of 8.4 allow for differing numbers of test results.

8.2.8 Collation of data and intermediate values

Specimens of recommended forms for the statistical analysis are given in Figure 2. Form A includes collated individual results, Form B contains calculated cell means, and Form C contains calculated standard deviations. For convenience, they are referred to simply as forms A, B and C (of Figure 2).

NOTE The use of forms A, B and C for statistical analysis is not a requirement of this document. Forms A, B, and C are, however, convenient for collation of intermediate values in manual calculation and can easily be adapted for use in spreadsheet software. Alternative layouts are often more appropriate for use with statistical software.

8.2.9 Original test results

See Form A of Figure 2, where

n_{ij} is the number of test results in the cell for laboratory i at level j;

y_{ijk} is any one of these test results (k = 1, 2, ..., n_{ij});

p_j is the number of laboratories reporting at least one test result for level j (after eliminating any test results designated as outliers or as erroneous).

8.2.10 Cell means (Form B of Figure 2)

These are derived from the results in Form A using Formula (2) as follows.

$$\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \tag{2}$$

The cell means should be recorded to at least one more significant digit than the reported test results (Form A).

8.2.11 Measures of cell spread (Form C of Figure 2)

These are derived from individual results (Form A, see 8.2.9) and means (Form B, see 8.2.10), as follows. For the general case, use the intracell standard deviation as in Formula (3),

$$s_{ij} = \sqrt{\frac{1}{n_{ij}-1} \sum_{k=1}^{n_{ij}} \left(y_{ijk} - \bar{y}_{ij} \right)^2} \tag{3}$$

or, equivalently

$$s_{ij} = \sqrt{\frac{1}{n_{ij}-1} \left[\sum_{k=1}^{n_{ij}} \left(y_{ijk} \right)^2 - \frac{1}{n_{ij}} \left[\sum_{k=1}^{n_{ij}} y_{ijk} \right]^2 \right]} \tag{4}$$

In using these formulae, care shall be taken to retain a sufficient number of digits in the calculations; i.e. every intermediate value shall be calculated to at least twice as many significant figures as required for the original reported data [see 7.3.2 b), 7.4.3 f)].

When using Form C, the standard deviation should be recorded to one more significant digit than the results in Form A. For values of n_{ij} less than 2, a dash should be inserted in Form C.

NOTE 1 If a cell ij contains two test results, the intracell standard deviation is

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

$$s_{ij} = \left| y_{ij1} - y_{ij2} \right| / \sqrt{2} \tag{5}$$

Therefore, for simplicity, absolute differences can be recorded in Form C and used for Cochran's test instead of standard deviations if all cells contain two test results.

NOTE 2 Formula (4) is prone to rounding error in computer calculation; for computer calculations, Formula (3) is more accurate.

8.2.12 Corrected or rejected data

As some of the data can be corrected or rejected on the basis of the tests mentioned in 8.3.3 and 8.3.4, the values of y_{ijk}, n_{ij} and p_j used for the final determinations of precision and mean can be different from the values referring to the original test results as recorded in forms A, B and C of Figure 2. Hence in reporting the final values for precision and trueness, it shall always be stated what data, if any, have been corrected or discarded.

8.3 Scrutiny of results for consistency and outliers

8.3.1 Approaches for scrutiny of data

From data collected on a number of specific levels, repeatability and reproducibility standard deviations are to be estimated. The presence of individual laboratories or values that appear to be inconsistent with all other laboratories or values can change the estimates, and decisions have to be made with respect to these values. Two approaches are introduced:

a) graphical consistency technique;

b) numerical outlier tests.

NOTE Reference [3] provides further discussion of the treatment of inter-laboratory data.

8.3.2 Graphical consistency technique

8.3.2.1 Two measures called Mandel's h and k statistics are used. As well as describing the variability of the measurement method, these help in laboratory evaluation.

8.3.2.2 Calculate the between-laboratory consistency statistic, h, for each laboratory by dividing the cell deviation (cell mean minus the grand mean for that level) by the standard deviation among cell means (for that level):

$$h_{ij} = \frac{\bar{y}_{ij} - \bar{\bar{y}}_j}{\sqrt{\frac{1}{p_j - 1} \sum_{i=1}^{p_j} \left(\bar{y}_{ij} - \bar{\bar{y}}_j \right)^2}} \tag{6}$$

in which, for $\bar{y}_{ij}$ see 8.2.10, and for $\bar{\bar{y}}_j$ see 8.4.4.

Plot the h_{ij} values for each cell in order of laboratory, in groups for each level (and separately grouped for the several levels examined by each laboratory) (see Figure C.7).

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

Form A – Recommended form for the collation of the original data [a]									
Laboratory	Level								
	1	2	...	...	j	...	...	$q-1$	q
1									
2									
⋮									
i	y_{i11} ⋮ y_{i1k} ⋮ y_{i1n}				y_{ij1} ⋮ y_{ijk} ⋮ y_{ijn}				y_{iq1} ⋮ y_{iqk} ⋮ y_{iqn}
⋮									
p									

Form B – Recommended form for the collation of the means [b]									
Laboratory	Level								
	1	2	...	...	j	...	...	$q-1$	q
1	$\overline{y}_{11}$								$\overline{y}_{1q}$
2									
⋮									
i					$\overline{y}_{ij}$				
⋮									
p	$\overline{y}_{p1}$								$\overline{y}_{pq}$

Form C – Recommended form for the collation of the measures of spread within cells [c]									
Laboratory	Level								
	1	2	...	...	j	...	...	$q-1$	q
1	s_{11}								s_{1q}
2									
⋮									
i					s_{ij}				
⋮									
p	s_{p1}								s_{pq}

[a] Form A contains collated individual results. Cells each contain all n results for the laboratory.

[b] Form B contains cell means as calculated in 8.2.10.

[c] Form C contains measures of spread (standard deviations) as calculated in 8.2.11.

Figure 2 — Recommended forms for the collation of results for analysis

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO 5725-2:2019

8.3.2.3 Calculate the within-laboratory consistency statistic, *k*, by first calculating the pooled within-cell standard deviation (Formula (7))

$$\sqrt{\frac{\sum s_{ij}^2}{p_j}} \tag{7}$$

for each level, and then calculate

$$k_{ij} = \frac{s_{ij}\sqrt{p_j}}{\sqrt{\sum s_{ij}^2}} \tag{8}$$

for each laboratory within each level.

Plot the k_{ij} values for each cell in order of laboratory, in groups for each level and separately grouped for the several levels examined by each laboratory (see Figure C.8).

8.3.2.4 Examination of the *h* and *k* plots can indicate that specific laboratories exhibit patterns of results that are markedly different from the others in the study. For example, laboratories can show consistently high or low within-cell variation and/or extreme cell means across many levels. If this occurs, the specific laboratory should be contacted to try to ascertain the cause of the discrepant behaviour. On the basis of the findings, the statistical expert can:

a) retain the laboratory's data for the moment;

b) ask the laboratory to redo the measurement (if feasible);

c) remove the laboratory's data from the study.

8.3.2.5 Various patterns can appear in the *h* plots. All laboratories can have both positive and negative *h* values at different levels of the experiment. Individual laboratories can tend to give either all positive or all negative *h* values, and the number of laboratories giving negative values is approximately equal to those giving positive values. Neither of these patterns is unusual or requires investigation, although the second of these patterns can suggest that a common source of laboratory bias exists. On the other hand, if all the *h* values for one laboratory are of one sign and the *h* values for the other laboratories are all of the other sign, then the reason should be sought. Likewise, if the *h* values for a laboratory are extreme and appear to depend on the experimental level in some systematic way, then the reason should be sought. Lines are drawn on the *h* plots corresponding to the indicators given in Table 7 and Table 8. These indicator lines serve as guides when examining patterns in the data.

8.3.2.6 If one laboratory stands out on the *k* plot as having many large values, then the reason should be sought; this indicates that it has a poorer repeatability than the other laboratories. A laboratory can give rise to consistently small *k* values because of such factors as excessive rounding of its data or an insensitive measurement scale. Lines are drawn on the *k* plots corresponding to the indicators given in (Table 7 and Table 8). These indicator lines serve as guides when examining patterns in the data.

8.3.2.7 When an *h* or *k* plot grouped by laboratory suggests that one laboratory has several *h* or *k* values near the indicator value line, the corresponding plot grouped by level should be studied. Often a value that appears large in a plot grouped by laboratory turns out to be reasonably consistent with other laboratories for the same level. If it is revealed as strongly different from values for the other laboratories, then the reason should be sought.

8.3.2.8 In addition to these *h* and *k* graphs, histograms of cell means and cell ranges can reveal the presence of, for example, two distinct populations. Such a case requires special treatment as the general underlying principle behind the methods described here assumes a single unimodal population.

© ISO 2019 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019

8.3.3 Numerical outlier technique

8.3.3.1 The following practice is recommended for dealing with outliers.

a) The tests recommended in 8.3.4 and 8.3.5 are applied to identify stragglers or outliers:

— if the test statistic is less than or equal to its 5 % critical value, the item tested is accepted as correct;

— if the test statistic is greater than its 5 % critical value and less than or equal to its 1 % critical value, the item tested is called a straggler and is indicated by a single asterisk;

— if the test statistic is greater than its 1 % critical value, the item is called a statistical outlier and is indicated by a double asterisk.

b) It is next investigated whether the stragglers and/or statistical outliers can be explained by some technical error, for example

— a mistake in performing the measurement,

— an error in computation,

— a simple clerical error in transcribing a test result, or

— analysis of the wrong sample.

Where the error was one of the computation or transcription type, the suspect result should be replaced by the correct value; where the error was from analysing a wrong sample, the result should be placed in its correct cell. After such correction has been made, the examination for stragglers or outliers should be repeated. If the explanation of the technical error is such that it proves impossible to replace the suspect test result, then it should be discarded as a "genuine" outlier that does not belong to the experiment proper.

c) When any stragglers and/or statistical outliers remain that have not been explained or rejected as belonging to an outlying laboratory, the stragglers are retained as correct items and the statistical outliers are discarded unless the statistician for good reason decides to retain them.

d) When the data for a cell have been rejected for Form B of Figure 2 under the above procedure, then the corresponding data shall be rejected for Form C of Figure 2, and vice versa.

8.3.3.2 The tests given in 8.3.4 and 8.3.5 are of two types. Cochran's test is a test of the within-laboratory variabilities and should be applied first, then any necessary action should be taken, with repeated tests if necessary. The other test (Grubbs') is primarily a test of between-laboratory variability, and can also be used (if $n > 2$) where Cochran's test has raised suspicions as to whether the high within-laboratory variation is attributable to only one of the test results in the cell.

8.3.4 Cochran's test

8.3.4.1 This document assumes that between laboratories only small differences exist in the within-laboratory variances. Experience, however, shows that this is not always the case, so that a test has been included to test the validity of this assumption. Several tests can be used for this purpose, but Cochran's test has been chosen.

© ISO 2019 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO 5725-2:2019