BRITISH STANDARD

BS ISO 5725-5:1998

Incorporating corrigendum August 2005

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

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

ICS 17.020

NO COPYING WITHOUT BSI PERMISSION EXCEPT AS PERMITTED BY COPYRIGHT LAW

AVA Documents Document Sharing Hub
British Standards

Full document details: BS ISO 5725-5:1998

National foreword

This British Standard is the UK implementation of ISO 5725-5:1998, incorporating corrigendum August 2005. It supersedes BS 5497-1:1987 which is withdrawn.

The UK participation in its preparation was entrusted to Technical Committee SS/6, Precision of test methods.

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.

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

This British Standard was published under the authority of the Standards Board and comes into effect on 15 October 1998

© BSI 2009

ISBN 978 0 580 63299 0

Amendments/corrigenda issued since publication

Date	Comments
31 August 2009	Implementation of ISO corrigendum August 2005 (equations (25) and (26) in 5.4.2 and equation (72) in 6.4.3 replaced)

Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

INTERNATIONAL STANDARD

ISO 5725-5

First edition
1998-07-15

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

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

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

Partie 5: Méthodes alternatives pour la détermination de la fidélité d'une méthode de mesure normalisée

Reference number
ISO 5725-5:1998(E)

Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

Contents

Page

1 Scope ... 1
2 Normative references ... 1
3 Definitions ... 2
4 Split-level design ... 2
4.1 Applications of the split-level design ... 2
4.2 Layout of the split-level design ... 2
4.3 Organization of a split-level experiment ... 3
4.4 Statistical model ... 4
4.5 Statistical analysis of the data from a split-level experiment ... 5
4.6 Scrutiny of the data for consistency and outliers ... 6
4.7 Reporting the results of a split-level experiment ... 7
4.8 Example 1: A split-level experiment — Determination of protein ... 7
5 A design for a heterogeneous material ... 13
5.1 Applications of the design for a heterogeneous material ... 13
5.2 Layout of the design for a heterogeneous material ... 14
5.3 Organization of an experiment with a heterogeneous material ... 15
5.4 Statistical model for an experiment with a heterogeneous material ... 16
5.5 Statistical analysis of the data from an experiment with a heterogeneous material ... 17
5.6 Scrutiny of the data for consistency and outliers ... 20
5.7 Reporting the results of an experiment on a heterogeneous material ... 21
5.8 Example 2: An experiment on a heterogeneous material ... 21
5.9 General formulae for calculations with the design for a heterogeneous material ... 29
5.10 Example 3: An application of the general formulae ... 30
6 Robust methods for data analysis ... 33
6.1 Applications of robust methods of data analysis ... 33
6.2 Robust analysis: Algorithm A ... 35
6.3 Robust analysis: Algorithm S ... 36
6.4 Formulae: Robust analysis for a particular level of a uniform-level design ... 38
6.5 Example 4: Robust analysis for a particular level of a uniform-level design ... 38
6.6 Formulae: Robust analysis for a particular level of a split-level design ... 42
6.7 Example 5: Robust analysis for a particular level of a split-level design ... 42
6.8 Formulae: Robust analysis for a particular level of an experiment on a heterogeneous material ... 45
6.9 Example 6: Robust analysis for a particular level of an experiment on a heterogeneous material ... 45

Annexes

A (normative) Symbols and abbreviations used in ISO 5725 ... 50
B (informative) Derivation of the factors used in algorithms A and S ... 53
C (informative) Derivation of equations used for robust analysis ... 55
D (informative) Bibliography ... 56

Descriptors: measurement, tests, test results, accuracy, precision, reproducibility, rules of calculation, statistical analysis.

Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

Foreword

ISO (the International Organization for Standardization) is a world-wide 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 organisations, 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.

ISO 5725-5 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

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

— *Part 1: General principles and definitions*

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

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

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

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

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

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

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

AVA Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

Introduction

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

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

0.3 ISO 5725-2 is concerned with estimating, by means of interlaboratory experiments, standard measures of precision, namely the repeatability standard deviation and the reproducibility standard deviation. It gives a basic method for doing this using the uniform-level design. This part of ISO 5725 describes alternative methods to this basic method.

a) With the basic method there is a risk that an operator may allow the result of a measurement on one sample to influence the result of a subsequent measurement on another sample of the same material, causing the estimates of the repeatability and reproducibility standard deviations to be biased. When this risk is considered to be serious, the split-level design described in this part of ISO 5725 may be preferred as it reduces this risk.

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

c) The basic method requires tests for outliers to be used to identify data that should be excluded from the calculation of the repeatability and reproducibility standard deviations. Excluding outliers can sometimes have a large effect on the estimates of repeatability and reproducibility standard deviations, but in practice, when applying the outlier tests, the data analyst may have to use judgement to decide which data to exclude. This part of ISO 5725 describes robust methods of data analysis that may be used to calculate repeatability and reproducibility standard deviations from data containing outliers without using tests for outliers to exclude data, so that the results are no longer affected by the data analyst's judgement.

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

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

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

1 Scope

This part of ISO 5725

— provides detailed descriptions of alternatives to the basic method for determining the repeatability and reproducibility standard deviations of a standard measurement method, namely the split-level design and a design for heterogeneous materials;

— describes the use of robust methods for analysing the results of precision experiments without using outlier tests to exclude data from the calculations, and in particular, the detailed use of one such method.

This part of ISO 5725 complements ISO 5725-2 by providing alternative designs that may be of more value in some situations than the basic design given in ISO 5725-2, and by providing a robust method of analysis that gives estimates of the repeatability and reproducibility standard deviations that are less dependent on the data analyst's judgement than those given by the methods described in ISO 5725-2.

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 3534-3:1985, *Statistics — Vocabulary and symbols — Part 3: Design of experiments.*

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.*

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

3 Definitions

For the purposes of this part of 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 Split-level design

4.1 Applications of the split-level design

4.1.1 The uniform level design described in ISO 5725-2 requires two or more identical samples of a material to be tested in each participating laboratory and at each level of the experiment. With this design there is a risk that an operator may allow the result of a measurement on one sample to influence the result of a subsequent measurement on another sample of the same material. If this happens, the results of the precision experiment will be distorted: estimates of the repeatability standard deviation σ_r will be decreased and estimates of the between-laboratory standard deviation σ_L will be increased. In the split-level design, each participating laboratory is provided with a sample of each of two similar materials, at each level of the experiment, and the operators are told that the samples are not identical, but they are not told by how much the materials differ. The split-level design thus provides a method of determining the repeatability and reproducibility standard deviations of a standard measurement method in a way that reduces the risk that a test result obtained on one sample will influence a test result on another sample in the experiment.

4.1.2 The data obtained at a level of a split-level experiment may be used to draw a graph in which the data for one material are plotted against the data for the other, similar, material. An example is given in figure 1. Such graphs can help identify those laboratories that have the largest biases relative to the other laboratories. This is useful when it is possible to investigate the causes of the largest laboratory biases with the aim of taking corrective action.

4.1.3 It is common for the repeatability and reproducibility standard deviations of a measurement method to depend on the level of the material. For example, when the test result is the proportion of an element obtained by chemical analysis, the repeatability and reproducibility standard deviations usually increase as the proportion of the element increases. It is necessary, for a split-level experiment, that the two similar materials used at a level of the experiment are so similar that they can be expected to give the same repeatability and reproducibility standard deviations. For the purposes of the split-level design, it is acceptable if the two materials used for a level of the experiment give almost the same level of measurement results, and nothing is to be gained by arranging that they differ substantially.

In many chemical analysis methods, the matrix containing the constituent of interest can influence the precision, so for a split-level experiment two materials with similar matrices are required at each level of the experiment. A sufficiently similar material can sometimes be prepared by spiking a material with a small addition of the constituent of interest. When the material is a natural or manufactured product, it can be difficult to find two products that are sufficiently similar for the purposes of a split-level experiment: a possible solution may be to use two batches of the same product. It should be remembered that the object of choosing the materials for the split-level design is to provide the operators with samples that they do not expect to be identical.

4.2 Layout of the split-level design

4.2.1 The layout of the split-level design is shown in table 1.

The p participating laboratories each test two samples at q levels.

The two samples within a level are denoted a and b, where a represents a sample of one material, and b represents a sample of the other, similar, material.

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

4.2.2 The data from a split-level experiment are represented by:

y_{ijk}

where

subscript i represents the laboratory (i = 1, 2, ..., p);

subscript j represents the level (j = 1, 2, ..., q);

subscript k represents the sample ($k = a$ or b).

4.3 Organization of a split-level experiment

4.3.1 Follow the guidance given in clause 6 of ISO 5725-1:1994 when planning a split-level experiment.

Subclause 6.3 of ISO 5725-1:1994 contains a number of formulae (involving a quantity denoted generally by A) that are used to help decide how many laboratories to include in the experiment. The corresponding formulae for the split-level experiment are set out below.

NOTE — These formulae have been derived by the method described in NOTE 24 of ISO 5725-1:1994.

To assess the uncertainties of the estimates of the repeatability and reproducibility standard deviations, calculate the following quantities.

For repeatability

$$A_r = 1{,}96\sqrt{1/[2(p-1)]} \tag{1}$$

For reproducibility

$$A_R = 1{,}96\sqrt{\left\{\left[1+2\left(\gamma^2-1\right)\right]^2+1\right\}\Big/\left[8\gamma^4(p-1)\right]} \tag{2}$$

with $\gamma = \sigma_R/\sigma_r$.

If the number n of replicates is taken as two in equations (9) and (10) of ISO 5725-1:1994, then it can be seen that equations (9) and (10) of ISO 5725-1:1994 are the same as equations (1) and (2) above, except that sometimes $p - 1$ appears here in place of p in ISO 5725-1:1994. This is a small difference, so table 1 and figures B.1 and B.2 of ISO 5725-1:1994 may be used to assess the uncertainty of the estimates of the repeatability and reproducibility standard deviations in a split-level experiment.

To assess the uncertainty of the estimate of the bias of the measurement method in a split-level experiment, calculate the quantity A as defined by equation (13) of ISO 5725-1:1994 with $n = 2$ (or use table 2 of ISO 5725-1:1994), and use this quantity as described in ISO 5725-1.

To assess the uncertainty of the estimate of a laboratory bias in a split-level experiment, calculate the quantity A_w as defined by equation (16) of ISO 5725-1:1994 with $n = 2$. Because the number of replicates in a split-level experiment is, in effect, this number of two, it is not possible to reduce the uncertainty of the estimate of laboratory bias by increasing the number of replicates. (If it is necessary to reduce this uncertainty, the uniform-level design should be used instead.)

4.3.2 Follow the guidance given in clauses 5 and 6 of ISO 5725-2:1994 with regard to the details of the organization of a split-level experiment. The number of replicates, n in ISO 5725-2, may be taken to be the number of split-levels in a split-level design, i.e. two.

The a samples should be allocated to the participants at random, and the b samples should also be allocated to the participants at random and in a separate randomization operation.

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

It is necessary in a split-level experiment for the statistical expert to be able to tell, when the data are reported, which result was obtained on material a and which on material b, at each level of the experiment. Label the samples so that this is possible, and be careful not to disclose this information to the participants.

Table 1 — Recommended form for the collation of data for the split-level design

Laboratory	Level									
	1		2			j			q	
	a	b	a	b		a	b		a	b
1										
2										
i										
p										

4.4 Statistical model

4.4.1 The basic model used in this part of ISO 5725 is given as equation (1) in clause 5 of ISO 5725-1:1994. It is stated there that for estimating the accuracy (trueness and precision) of a measurement method, it is useful to assume that every measurement result is the sum of three components:

$$y_{ijk} = m_j + B_{ij} + e_{ijk} \tag{3}$$

where, for the particular material tested,

m_j represents the general average (expectation) at a particular level $j = 1, ..., q$;

B_{ij} represents the laboratory component of bias under repeatability conditions in a particular laboratory $i = 1, ..., p$ at a particular level $j = 1, ..., q$;

e_{ijk} represents the random error of test result $k = 1, ..., n$, obtained in laboratory i at level j, under repeatability conditions.

4.4.2 For a split-level experiment, this model becomes:

$$y_{ijk} = m_{jk} + B_{ij} + e_{ijk} \tag{4}$$

This differs from equation (3) in 4.4.1 in only one feature: the subscript k in m_{jk} implies that according to equation (4) the general average may now depend on the material a or b (k = 1 or 2) within the level j.

The lack of a subscript k in B_{ij} implies that it is assumed that the bias associated with a laboratory i does not depend on the material a or b within a level. This is why it is important that the two materials should be similar.

4.4.3 Define the cell averages as:

$$y_{ij} = (y_{ija} + y_{ijb}) / 2 \tag{5}$$

and the cell differences as:

$$D_{ij} = y_{ija} - y_{ijb} \tag{6}$$

AVA Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

4.4.4 The general average for a level j of a split-level experiment may be defined as:

$$m_j = (m_{ja} + m_{jb}) / 2 \tag{7}$$

4.5 Statistical analysis of the data from a split-level experiment

4.5.1 Assemble the data into a table as shown in table 1. Each combination of a laboratory and a level gives a "cell" in this table, containing two items of data, y_{ija} and y_{ijb}.

Calculate the cell differences D_{ij} and enter them into a table as shown in table 2. The method of analysis requires each difference to be calculated in the same sense

$$a - b$$

and the sign of the difference to be retained.

Calculate the cell averages y_{ij} and enter them into a table as shown in table 3.

4.5.2 If a cell in table 1 does not contain two test results (for example, because samples have been spoilt, or data have been excluded following the application of the outlier tests described later) then the corresponding cells in tables 2 and 3 both remain empty.

4.5.3 For each level j of the experiment, calculate the average D_j and standard deviation s_{Dj} of the differences in column j of table 2:

$$D_j = \sum D_{ij} / p \tag{8}$$

$$s_{Dj} = \sqrt{\sum \left(D_{ij} - D_j\right)^2 / (p-1)} \tag{9}$$

Here, Σ represents summation over the laboratories $i = 1, 2, ..., p$.

If there are empty cells in table 2, p is now the number of cells in column j of table 2 containing data and the summation is performed over non-empty cells.

4.5.4 For each level j of the experiment, calculate the average y_j and standard deviation s_{yj} of the averages in column j of table 3, using:

$$y_j = \sum y_{ij} / p \tag{10}$$

$$s_{yj} = \sqrt{\sum \left(y_{ij} - y_j\right)^2 / (p-1)} \tag{11}$$

Here, Σ represents summation over the laboratories $i = 1, 2, ..., p$.

If there are empty cells in table 3, p is now the number of cells in column j of table 3 containing data and the summation is performed over non-empty cells.

4.5.5 Use tables 2 and 3 and the statistics calculated in 4.5.3 and 4.5.4 to examine the data for consistency and outliers, as described in 4.6. If data are rejected, recalculate the statistics.

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

4.5.6 Calculate the repeatability standard deviation s_{rj} and the reproducibility standard deviation s_{Rj} from:

$$s_{rj} = s_{Dj} / \sqrt{2} \tag{12}$$

$$s_{Rj}^2 = s_{yj}^2 + s_{rj}^2 / 2 \tag{13}$$

4.5.7 Investigate whether s_{rj} and s_{Rj} depend on the average y_j, and, if so, determine the functional relationships, using the methods described in subclause 7.5 of ISO 5725-2:1994.

Table 2 — Recommended form for tabulation of cell differences for the split-level design

Laboratory	Level					
	1	2		j		q
1						
2						
i						
p						

Table 3 — Recommended form for tabulation of cell averages for the split-level design

Laboratory	Level					
	1	2		j		q
1						
2						
i						
p						

4.6 Scrutiny of the data for consistency and outliers

4.6.1 Examine the data for consistency using the h statistics, described in subclause 7.3.1 of ISO 5725-2:1994.

To check the consistency of the cell differences, calculate the h statistics as:

$$h_{ij} = \left(D_{ij} - D_j\right) / s_{Dj} \tag{14}$$

AVA Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

To check the consistency of the cell averages, calculate the h statistics as:

$$h_{ij} = \left(y_{ij} - y_j\right) / s_{yj} \tag{15}$$

To show up inconsistent laboratories, plot both sets of these statistics in the order of the levels, but grouped by laboratory, as shown in figures 2 and 3. The interpretation of these graphs is discussed fully in subclause 7.3.1 of ISO 5725-2:1994. If a laboratory is achieving generally worse repeatability than the others, then it will show up as having an unusually large number of large h statistics in the graph derived from the cell differences. If a laboratory is achieving results that are generally biased, then it will show up as having h statistics mostly in one direction on the graph derived from the cell averages. In either case, the laboratory should be asked to investigate and report their findings back to the organizer of the experiment.

4.6.2 Examine the data for stragglers and outliers using Grubbs' tests, described in subclause 7.3.4 of ISO 5725-2:1994.

To test for stragglers and outliers in the cell differences, apply Grubbs' tests to the values in each column of table 2 in turn.

To test for stragglers and outliers in the cell averages, apply Grubbs' tests to the values in each column of table 3 in turn.

The interpretation of these tests is discussed fully in subclause 7.3.2 of ISO 5725-2:1994. They are used to identify results that are so inconsistent with the remainder of the data reported in the experiment that their inclusion in the calculation of the repeatability and reproducibility standard deviations would affect the values of these statistics substantially. Usually, data shown to be outliers are excluded from the calculations, and data shown to be stragglers are included, unless there is a good reason for doing otherwise. If the tests show that a value in one of tables 2 or 3 is to be excluded from the calculation of the repeatability and reproducibility standard deviations, then the corresponding value in the other of these tables should also be excluded from the calculation.

4.7 Reporting the results of a split-level experiment

4.7.1 Advice is given in subclause 7.7 of ISO 5725-2:1994 on:

— reporting the results of the statistical analysis to the panel;

— decisions to be made by the panel; and

— the preparation of a full report.

4.7.2 Recommendations on the form of a published statement of the repeatability and reproducibility standard deviations of a standard measurement method are given in subclause 7.1 of ISO 5725-1:1994.

4.8 Example 1: A split-level experiment — Determination of protein

4.8.1 Table 4 contains the data from an experiment [5] which involved the determination by combustion of the content of protein in feeds. There were nine participating laboratories, and the experiment contained 14 levels. Within each level, two feeds were used having similar mass fraction of protein in feed.

4.8.2 Tables 5 and 6 show the cell averages and differences, calculated as described in clause 4.5.1, for just Level 14 (j = 14) of the experiment.

Using equations (8) and (9) in 4.5.3, the differences in table 5 give:

D_{14} = 8,34 %

s_{D14} = 0,436 1 %

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

and applying equations (10) and (11) in 4.5.4 to the averages in table 6 gives:

y_{14} = 85,46 %

s_{Y14} = 0,453 4 %

so the repeatability and reproducibility standard deviations are, using equations (12) and (13) in 4.5.6:

s_{r14} = 0,31 %

s_{R14} = 0,50 %

Table 7 gives the results of the calculations for the other levels.

4.8.3 Figure 1 shows the results for samples a from table 4 plotted against the corresponding results for samples b, for Level 14, in a "Youden plot". Laboratory 5 gives a point in the bottom left-hand corner of the graph, and Laboratory 1 gives a point in the top right-hand corner: this indicates that the data from Laboratory 5 have a consistent negative bias over samples a and b, and that the data from Laboratory 1 have a consistent positive bias over the two samples. It is common to find this sort of pattern when plotting the data from a split-level design as in figure 1. The figure also shows that the results for Laboratory 4 are unusual, as the point for this laboratory is some distance from the line of equality for the two samples. The other laboratories form a group in the middle of the plot. This figure thus provides a case for investigating the causes of the biases at the three laboratories.

NOTE — For further information on the interpretation of "Youden plots", see references [7] and [8].

4.8.4 The values of the h statistics, calculated as described in 4.6.1, are shown in tables 5 and 6, for only Level 14. The values for all levels are plotted in figures 2 and 3.

In figure 3, the h statistics for cell averages show that Laboratory 5 gave negative h statistics at all levels, indicating a consistent negative bias in their data. In the same figure, Laboratories 8 and 9 gave h statistics that are nearly all positive, indicating consistent positive biases in their data (but smaller than the negative bias in Laboratory 5). Also, the h statistics for Laboratories 1, 2 and 6 indicate a bias that changes with level in each of these laboratories. Such interactions between the laboratories and the levels may provide clues as to the causes of the laboratory biases.

Figure 2 does not reveal any noteworthy pattern.

4.8.5 Values of the Grubbs' statistics are given in table 8. These tests again indicate that the data from Laboratory 5 are suspect.

4.8.6 At this point in the analysis, the statistical expert should initiate an investigation at Laboratory 5 of the possible causes of the suspect data, before proceeding with the analysis of the data. If the cause cannot be identified, there is a case in this instance for excluding all data from Laboratory 5 from the calculation of the repeatability and reproducibility standard deviations. The analysis would then continue with an investigation of possible functional relationships between the repeatability and reproducibility standard deviations and the general average. This does not raise any issues that have not already been covered in ISO 5725-2, so it will not be considered here.

Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

Table 4 — Example 1: Determination of mass fraction of protein in feed, expressed as a percentage

Laboratory	Level									
	1		2		3		4		5	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	11,11	10,34	10,91	9,81	13,74	13,48	13,79	13,00	15,89	15,26
2	11,12	9,94	11,38	10,31	14,00	13,12	13,44	13,06	15,69	15,10
3	11,26	10,46	10,95	10,51	13,38	12,70	13,54	13,18	15,83	15,73
4	11,07	10,41	11,66	9,95	13,01	13,16	13,58	12,88	15,08	15,63
5	10,69	10,31	10,98	10,13	13,24	13,33	13,32	12,59	15,02	14,90
6	11,73	11,01	12,31	10,92	14,01	13,66	14,04	13,64	16,43	15,94
7	11,13	10,36	11,38	10,44	12,94	12,44	13,63	13,06	15,75	15,56
8	11,21	10,51	11,32	10,84	13,09	13,76	13,85	13,49	15,98	15,89
9	11,80	11,21	11,35	9,88	13,85	14,46	13,96	13,77	16,51	15,72
Laboratory	**Level**									
	6		7		8		9		10	
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	20,14	19,78	20,33	20,06	46,45	44,42	52,05	49,40	65,84	59,14
2	19,25	20,25	20,36	19,94	46,69	44,62	51,94	48,81	66,31	59,19
3	20,48	19,86	20,56	20,11	46,90	44,56	52,18	48,90	66,06	58,52
4	21,54	20,06	20,64	20,46	47,13	45,29	51,73	48,56	65,93	58,93
5	19,90	19,66	20,56	19,24	45,83	43,73	50,84	47,91	64,19	57,94
6	20,31	20,27	20,85	20,63	46,86	43,96	52,18	49,03	65,73	58,77
7	20,00	20,56	20,25	20,19	46,25	44,31	52,25	49,44	66,06	59,19
8	20,43	20,69	20,85	20,27	47,11	44,40	52,44	48,81	65,66	59,38
9	20,64	21,01	20,78	20,89	47,09	45,15	52,19	48,46	66,33	59,47
Laboratory	**Level**									
	11		12		13		14			
	a	*b*	*a*	*b*	*a*	*b*	*a*	*b*		
1	84,16	80,86	85,38	81,71	87,64	88,23	90,24	82,10		
2	84,50	81,06	85,56	82,44	88,81	88,38	89,88	81,44		
3	82,26	79,43	85,26	82,15	88,58	88,12	89,48	81,67		
4	84,39	80,08	85,20	81,76	88,47	87,98	90,04	80,73		
5	81,71	79,01	83,58	79,74	86,43	86,19	88,59	80,46		
6	82,85	81,16	84,44	80,90	87,78	86,89	89,40	80,88		
7	86,25	81,00	84,88	81,44	88,06	88,00	89,31	81,38		
8	84,59	81,16	84,96	81,71	88,50	87,98	89,94	81,56		
9	83,05	80,93	84,73	81,94	88,24	88,05	89,75	81,35		

AVA Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

Table 5 — Example 1: Cell differences for Level 14

Laboratory	Cell difference %	h statistic
1	8,14	– 0,459
2	8,44	0,229
3	7,81	– 1,215
4	9,31	2,224
5	8,13	– 0,482
6	8,52	0,413
7	7,93	– 0,940
8	8,38	0,092
9	8,40	0,138

Table 6 — Example 1: Cell averages for Level 14

Laboratory	Cell difference %	h statistic
1	86,170	1,576
2	85,660	0,451
3	85,575	0,263
4	85,385	– 0,156
5	84,525	– 2,052
6	85,140	– 0,696
7	85,345	– 0,244
8	85,750	0,649
9	85,550	0,208

Table 7 — Example 1: Values of averages, average differences, and standard deviations calculated from the data for all 14 levels in table 4

Level	Number of laboratories	General average	Average difference	Standard deviations			
j	p	y_j %	D_j %	s_{yj} %	s_{Dj} %	s_{rj} %	s_{Rj} %
1	9	10,87	0,73	0,35	0,21	0,15	0,36
2	9	10,84	1,05	0,36	0,43	0,30	0,42
3	9	13,41	0,13	0,44	0,55	0,39	0,52
4	9	13,43	0,50	0,30	0,21	0,15	0,32
5	9	15,66	0,27	0,39	0,40	0,29	0,44
6	9	20,27	0,06	0,40	0,73	0,52	0,54
7	9	20,39	0,38	0,30	0,41	0,29	0,37
8	9	45,60	2,21	0,44	0,37	0,26	0,47
9	9	50,40	3,16	0,44	0,35	0,25	0,47
10	9	62,37	6,84	0,53	0,40	0,28	0,57
11	9	82,14	3,23	1,01	1,08	0,77	1,15
12	9	83,17	3,45	0,74	0,46	0,33	0,77
13	9	87,91	0,30	0,69	0,41	0,29	0,72
14	9	85,46	8,34	0,45	0,44	0,31	0,50

AVA Documents
Document Sharing Hub

Full document details: BS ISO 5725-5:1998

Table 8 — Example 1: Values of Grubbs' statistics

Grubbs' statistics for differences				
Level	**One smallest**	**Two smallest**	**Two largest**	**One largest**
1	1,653	0,5081	0,3139	2,125
2	1,418	0,3945	0,4738	1,535
3	1,462	0,3628	0,5323	1,379
4	1,490	0,5841	0,4771	1,414
5	2,033	0,3485	0,6075	1,289
6	1,456	0,5490	0,3210	1,947
7	1,185	0,6820	0,1712	2,296* (5)
8	0,996	0,7571	0,1418* (6; 8)	1,876
9	1,458	0,5002	0,3092	1,602
10	1,474	0,3360	0,4578	1,737
11	1,422	0,5089	0,2943	1,865
12	1,418	0,6009	0,2899	1,956
13	2,172	0,2325	0,6326	1,444
14	1,215	0,6220	0,2362	2,224* (4)
Grubbs' statistics for cell averages				
Level	**One smallest**	**Two smallest**	**Two largest**	**One largest**
1	1,070	0,6607	0,1291* (6; 9)	1,832
2	1,318	0,6288	0,2118	2,165
3	1,621	0,4771	0,4077	1,680
4	1,591	0,5339	0,3807	1,429
5	1,794	0,4018	0,5009	1,333
6	1,291	0,4947	0,4095	1,386
7	1,599	0,5036	0,4391	1,470
8	1,872	0,3753	0,4536	1,404
9	2,328* (5)	0,1317* (4; 5)	0,7417	1,025
10	2,456** (5)	-	-	1,000
11	1,756	0,2469	0,5759	1,472
12	2,037	0,1063* (5; 6)	0,7116	1,130
13	2,308* (5)	0,0733** (5; 6)	0,7777	0,994
14	2,052	0,2781	0,5486	1,576

NOTE — Numbers in brackets indicate the laboratories that give rise to the stragglers or outliers.

The critical values of the Grubbs' test statistics for 9 laboratories, whether applied to the differences or the cell averages, are as follows.

	Straggler (*)	Outlier (**)
Grubbs' test for a single outlier	2,215	2,387
Grubbs' test for a pair of outliers	0,149 2	0,085 1

AVA Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

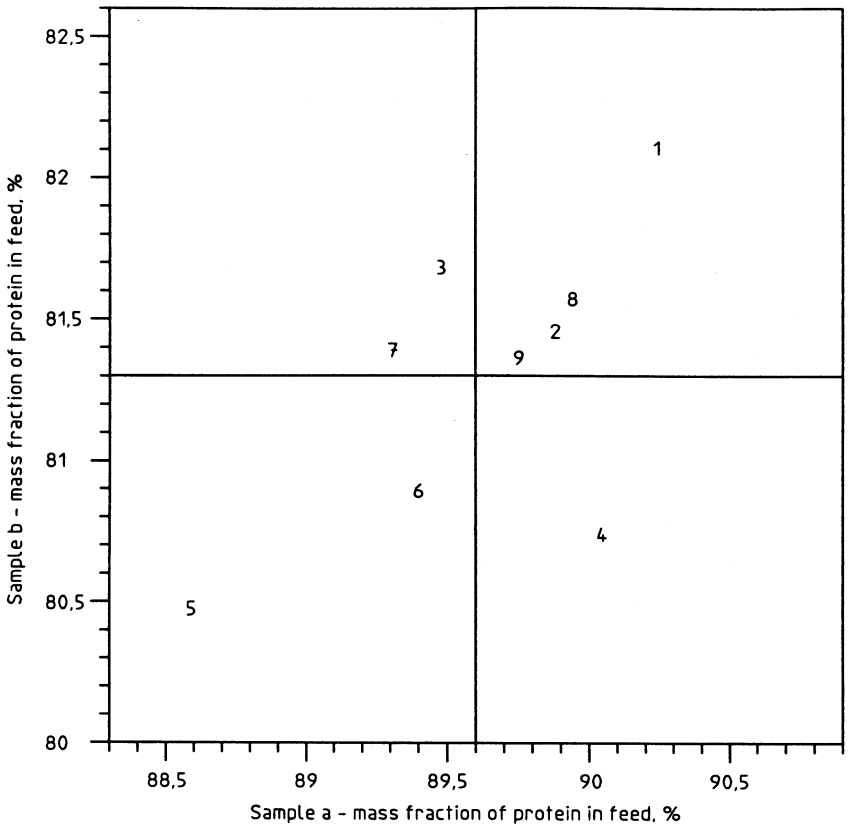

Figure 1 — Example 1: Data obtained at Level 14

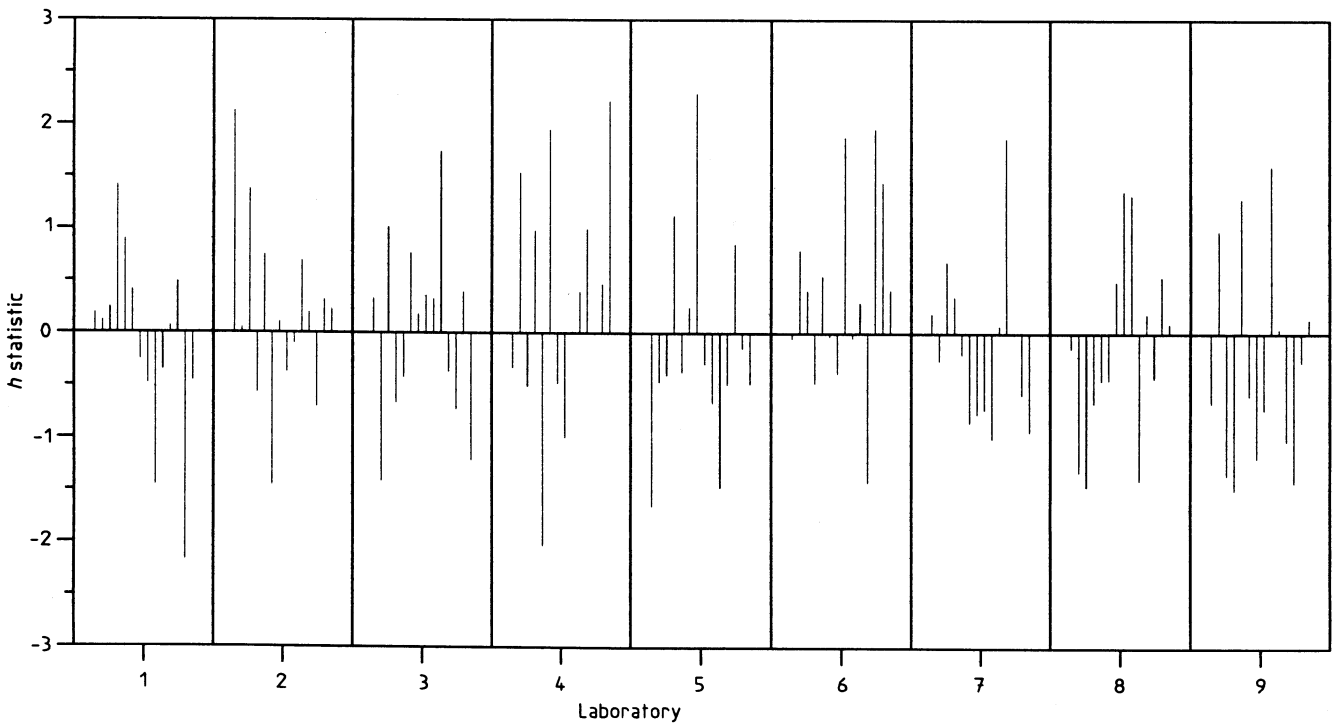

Figure 2 — Example 1: Consistency check on cell differences (grouped by laboratory)

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

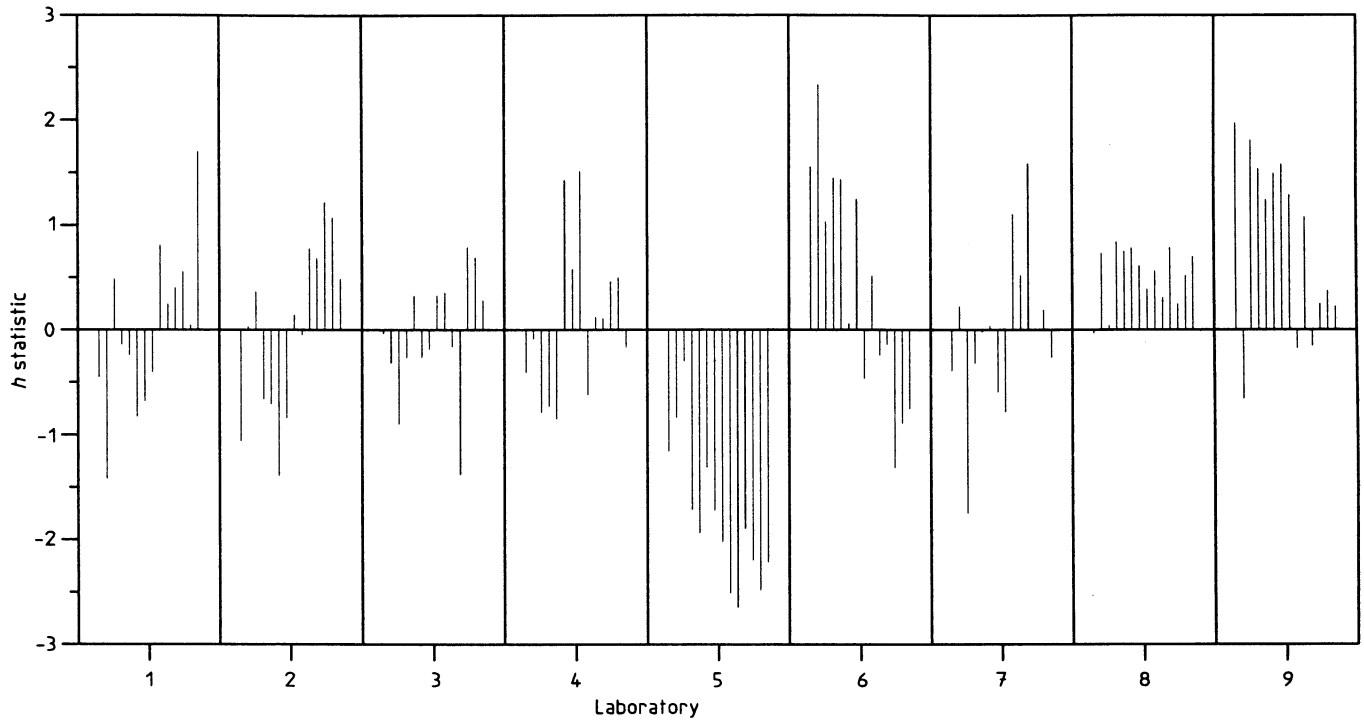

Figure 3 — Example 1: Consistency check on cell averages (grouped by laboratory)

5 A design for a heterogeneous material

5.1 Applications of the design for a heterogeneous material

5.1.1 An example of a heterogeneous material is leather; no two hides are the same, and the properties of leather vary substantially within a hide. A common test that is applied to leather is a tensile strength test in BS 3144 [3]. This is performed on dumbell-shaped specimens (BS 3144 specifies the number of such specimens to cut from a hide, and also their position and orientation within the hide, so the natural definition of a "sample" to use when testing leather is a complete hide). If a precision experiment is performed using the uniform level design described in ISO 5725-2, in which each laboratory is sent one hide at each level of the experiment, and two test results are obtained on each hide, variation between hides will add to the between-laboratory variation, and so increase the reproducibility standard deviation. However, if each laboratory is sent two hides at each level, and two test results are obtained on each hide, the data can be used to estimate the variation between hides and to calculate a value for the reproducibility standard deviation of the test method from which the variation between-hides has been removed.

5.1.2 Another example of a heterogeneous material is sand (that might be used, for example, for making concrete). This is laid down, by the action of wind or water, in strata that always contain graduations in particle size, so when sand is used the particle size distribution is always of interest. In concrete technology the particle size distribution of sand is measured by sieve testing (e.g. BS 812-103 [1]). To test a sand, a bulk sample is taken from the product, then one or more test portions are produced from the bulk sample. Typically, the bulk sample will be about 10 kg in mass, and the test portions will be about 200 g. Because of the natural variability of the material, there will always be some variability between bulk samples of the same product. Hence, as with leather, if a uniform level experiment is performed in which each laboratory is sent one bulk sample at each level, the variability between the bulk samples will increase the calculated reproducibility standard deviation of the test method, but if laboratories are sent two bulk samples at each level, then values for the reproducibility standard deviation can be calculated that exclude this variation.

5.1.3 The above examples also highlight another characteristic of heterogeneous materials: because of the variability of the material, the specimen or test portion preparation can be an important source of variation. Thus with leather, the process of cutting specimens from a hide can have a large influence on the measured tensile

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998

strength, and with sieve tests on sand the process of preparing test portions from bulk samples is usually the major source of variability in the test method. If specimens or test portions are prepared for a precision experiment in a way that does not correspond to normal practice (in an attempt to produce identical "samples") then the values of repeatability and reproducibility standard deviations produced by the experiment will not be representative of the variability experienced in practice. There are situations in which it can be desirable to produce identical "samples" by some special process designed to eliminate, as far as possible, the variability of the material (for example, for a proficiency test, or when a precision experiment is used as part of a programme of work during the development of a measurement method). However, when the aim of the precision experiment is to discover the variability that will be experienced in practice (for example, when vendors and purchasers test samples of the same product) then it is necessary for the variability arising as a consequence of the heterogeneity of the material to be included in the measures of the precision of the measurement method.

Care should also be taken to ensure that each test result in an experiment is obtained by carrying out the test procedure independently of other tests. This will not be so if some stages of the specimen preparation are shared by several specimens, so that a bias or deviation introduced by the preparation will have a common influence on the test results derived from these specimens.

5.1.4 The design for heterogeneous materials proposed in this clause yields information about the variability between samples that is not obtainable from the uniform level design described in ISO 5725-2. There is, inevitably, a cost associated with obtaining extra information: the proposed design requires more samples to be tested. This extra information may be valuable. In the leather example discussed in 5.1.1, information about the variability between hides could be used to decide how many hides to test when assessing the quality of a consignment, or to decide between testing more hides with fewer specimens per hide or fewer hides with more specimens per hide. In the sand example discussed in 5.1.2, information about the variability between bulk samples could be used to decide if the procedure for taking bulk samples is satisfactory or in need of improvement.

5.1.5 The design described in this clause is applicable to experiments involving three factors arranged in a hierarchy: with a factor "laboratories" at the highest level in the hierarchy, a factor "samples within laboratories" as the next level in the hierarchy, and a factor "test results within samples" as the lowest level of the hierarchy. Another case that can be encountered in practice is of a three-factor hierarchy with "laboratories" at the highest level, "test results within laboratories" as the next level, and "determinations within test results" as the lowest level. This would arise if the participating laboratories in a precision experiment were each sent a single sample of a homogeneous material, were asked to carry out two (or perhaps more) tests per sample, and if each test involved a number of determinations and the test result is calculated as the average of the determinations. The formulae given in 5.5, 5.6 and 5.9 may be applied to data obtained in such an experiment, but the repeatability and reproducibility standard deviations have to be calculated in a slightly different manner to that given here (see NOTE 2 to 5.5.5). It is also necessary to specify the number of determinations that are to be averaged to give a test result, because this affects the values of the repeatability and reproducibility standard deviations.

5.2 Layout of the design for a heterogeneous material

5.2.1 The layout of the design for a heterogeneous material is shown in table 9.

The p participating laboratories are each provided with two samples at q levels, and obtain two test results on each sample. Thus each cell in the experiment contains four test results (two test results for each of two samples).

It is possible to generalize this simple design, by allowing for more than two samples per laboratory per level, or for more than two test results per sample. The calculations required by the more general design are much more complicated than those required by the case with two test results per sample and two samples per laboratory per level. However, the principles of the more general design are the same as for the simple design, so the calculations will be set out in detail here for the simple design. Formulae for calculating values for the repeatability and reproducibility standard deviations for the general design are given below in 5.9, and an example of their application in 5.10.

5.2.2 The data from the design for a heterogeneous material are represented by:

y_{ijtk}

Documents
Document Sharing Hub
Full document details: BS ISO 5725-5:1998