TECHNICAL REPORT

ISO/TR 22971

First edition
2005-01-15

Accuracy (trueness and precision) of measurement methods and results — Practical guidance for the use of ISO 5725-2:1994 in designing, implementing and statistically analysing interlaboratory repeatability and reproducibility results

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f4aeadc831/iso-tr-22971-2005

Exactitude (justesse et fidélité) des résultats et méthodes de mesure — Lignes directrices pratiques pour l'utilisation de l'ISO 5725-2:1994 pour la conception, la mise en œuvre et l'analyse statistique des résultats de répétabilité et de reproductibilité interlaboratoires

Reference number
ISO/TR 22971:2005(E)

© ISO 2005

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f4aeadc831/iso-tr-22971-2005

© ISO 2005

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

 © ISO 2005 – All rights reserved

Contents

Page

Foreword v

Introduction vii

1 Scope 1

2 Organization of an inter-laboratory programme 1
2.1 Requirements for a precision experiment 1
2.2 The responsibilities of the personnel involved in a precision experiment 1

3 Critical examination of the data 3
3.1 Description of the data 3
3.2 Tests for outliers 7
3.3 Conclusions 9

4 Estimation of repeatability and reproducibility standard deviations 11
4.1 Analysis of variance 11
4.2 Description of the model 11
4.3 Examples 13
4.4 Use of repeatability and reproducibility limits 15

5 Worked examples using statistical software 16
5.1 General 16
5.2 Determination of sulfur content in coal 16
5.3 Thermometric titration of creosote oil 23

Annex A (normative) **Symbols and abbreviations** 30

Bibliography 33

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f4acadc831/iso-tr-22971-2005

© ISO 2005 – All rights reserved

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. 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.

In exceptional circumstances, when a technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example), it may decide by a simple majority vote of its participating members to publish a Technical Report. A Technical Report is entirely informative in nature and does not have to be reviewed until the data it provides are considered to be no longer valid or useful.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

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.

ISO/TR 22971 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f4aeadc831/iso-tr-22971-2005

© ISO 2005 – All rights reserved

Introduction

ISO 5725 consists of six parts, the general structure of which is shown in Figure 1.

ISO 5725-2 was developed as a guidance document for ISO Technical Committees and other organizations responsible for undertaking inter-laboratory studies for characterizing the variability of standard measurement methods. Two measures of variability, repeatability and reproducibility, are accepted in many disciplines as representative of data encountered in measurement processes.

Repeatability refers to the variability among measurements made on nominally identical samples or materials under identical circumstances. It is recognized that, because of unknown or uncontrollable factors which influence the measurement process, repeated measurements will usually not agree. The extent of this variability can be expressed by a standard deviation, called the repeatability standard deviation, of the results of within-laboratory comparisons.

Reproducibility refers to the variability among measurements made on identical samples or materials under differing conditions by different laboratories following the same standard measurement method. Reproducibility includes effects caused by differences among instruments, reagents, operators, laboratories, and environmental conditions. The variability of results under these conditions may be described by a standard deviation called the reproducibility standard deviation.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

This guidance document is divided into four clauses in addition to the Scope (Clause 1):

— Clause 2, Organization of an inter-laboratory programme, deals with the organization of the inter-laboratory test and covers the roles of the executive officer, laboratory personnel, and statistician in preparing for and administering the test; the choice of materials and levels of interest for the test; and the selection of laboratories. It also describes how the number of replicate measurements (to be made on each sample) is to be statistically treated and the manner in which the resulting data are to be reported.

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-33e1-4267-ad42-16f1aeadc831/iso-tr-22971-2005

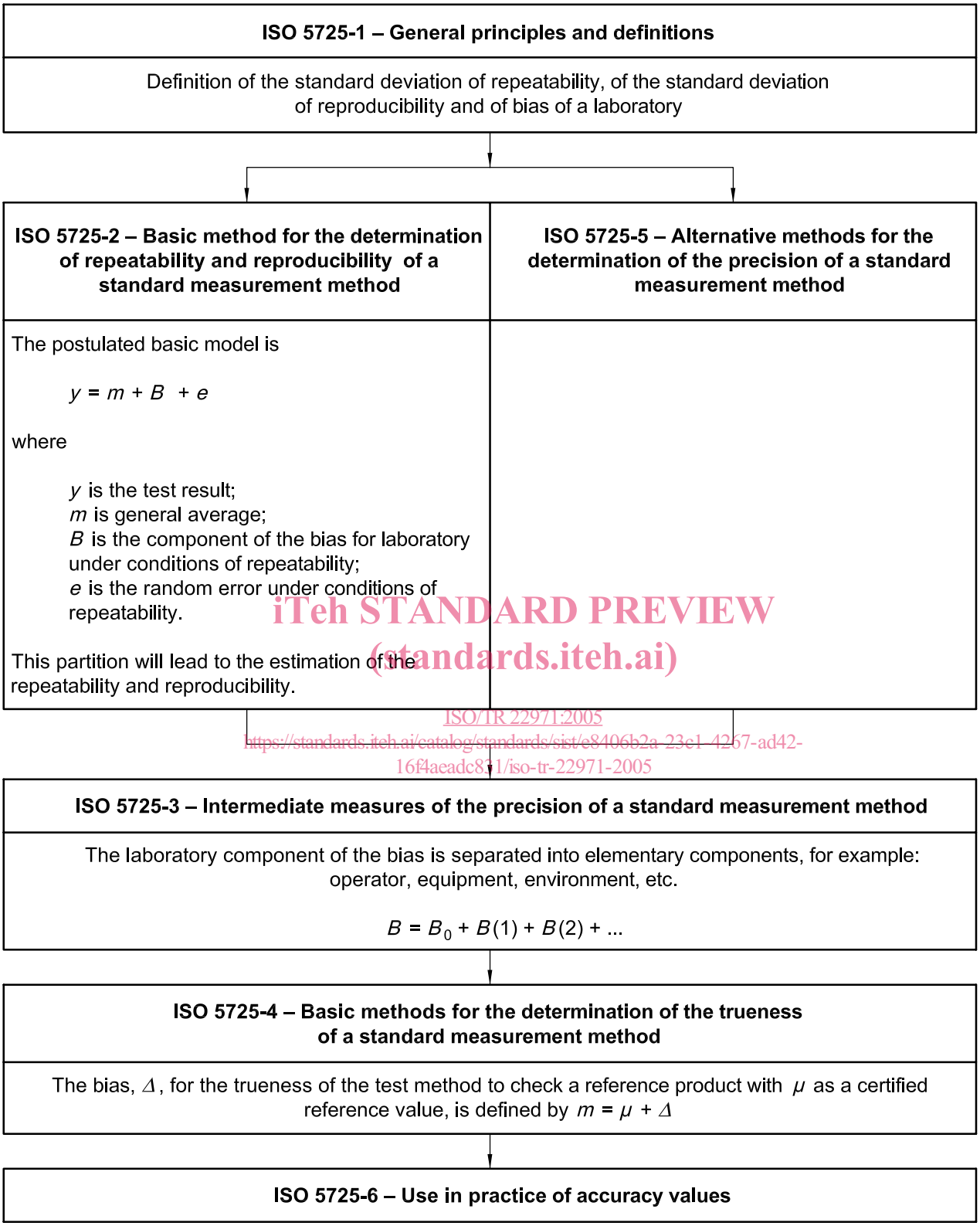
— Clause 3, Critical examination of the data, deals with data using graphical and numerical procedures. Guidance is given as to when data are anomalous, i.e. if they are inconsistent with other data from the study, and for outlier tests that are used to identify the presence or absence of anomalous data.

— Clause 4, Estimation of repeatability and reproducibility standard deviations, deals with the estimation and interpretation of repeatability and reproducibility standard deviations. Also included is a comparison of the relative contributions of the repeatability and reproducibility standard deviations to the total variability of the test method.

— Clause 5, Worked examples using statistical software, deals with worked examples that highlight various techniques that can be used.

It is recommended that this guidance document be read in conjunction with ISO 5725-2 and should not be used as a replacement for ISO 5725-2.

© ISO 2005 – All rights reserved



iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23e1-4267-ad42-16f4aeadc871/iso-tr-22971-2005

Figure 1 — Structure of ISO 5725 — Application of a standardized test method to the analyses of a sample or product in different laboratories

© ISO 2005 – All rights reserved

Accuracy (trueness and precision) of measurement methods and results — Practical guidance for the use of ISO 5725-2:1994 in designing, implementing and statistically analysing interlaboratory repeatability and reproducibility results

1 Scope

This Technical Report provides users with practical guidance to the use of ISO 5725-2:1994 and presents simplified step-by-step procedures for the design, implementation, and statistical analysis of inter-laboratory studies for assessing the variability of a standard measurement method and on the determination of repeatability and reproducibility of data obtained in inter-laboratory testing.

2 Organization of an inter-laboratory programme

2.1 Requirements for a precision experiment

iTeh STANDARD PREVIEW
(standards.iteh.ai)

The whole experiment is organized

a) to provide a complete set of results for:

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-

— p_{Lab} number of laboratories demonstrating that the test procedure is well controlled and for quantifying the observed scatter, estimated by the reproducibility;

NOTE The symbol p_{Lab} used in this Technical Report has the same meaning as the symbol p used in ISO 5725-2:1994. The change was made to clearly distinguish this symbol from the symbol, P, used for "probability". The lowercase and uppercase P's are sometimes difficult to distinguish, particularly as subscripts.

— q number of samples or products representing different levels of results or performance. A minimum value for q is two, but from five to ten is more appropriate for demonstrating that the test procedure is able to discriminate correctly between levels;

— n number of replications cell demonstrating that the test procedure is well controlled within a single laboratory. When the number of laboratories and of levels is sufficient, at least two determinations are required;

b) to analyse statistically (see Clauses 2 and 3) a table of results reported by p_{Lab} laboratories analysing q samples, tested n times under conditions of repeatability.

The table of the results submitted to the executive officer is shown in Table 1 (see ISO 5725-2:1994, 7.2.8).

2.2 The responsibilities of the personnel involved in a precision experiment

2.2.1 General

An inter-laboratory programme is very expensive, both in terms of its co-ordination and its participation. Hence, the performance testing should be well co-ordinated and planned. In any inter-laboratory programme, it is necessary to consider three types of activity as shown in Figure 2.

© ISO 2005 – All rights reserved

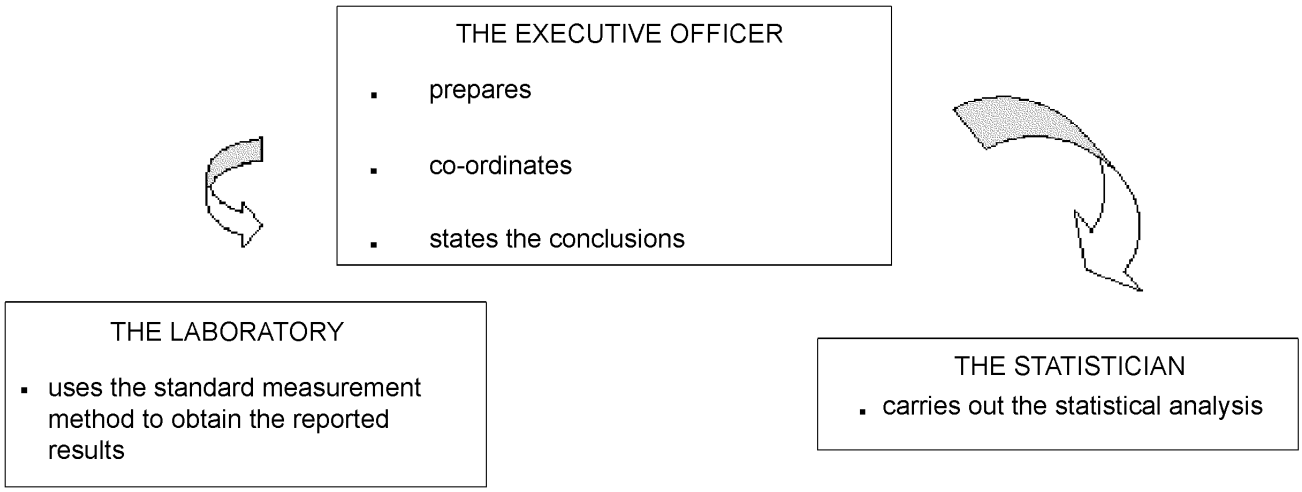


Figure 2 — The responsibilities of executive functions

2.2.2 Executive officer

The executive officer's main tasks are

— to organize the inter-laboratory programme, with the advice of the statistician for the construction of the experimental design;

— to co-ordinate the progress;

— to state the conclusions.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f1aeadc831/iso-tr-22971-2005

The duties of the executive officer may be undertaken by more than one person. However, only one person should be responsible for the entire programme. The executive officer should be familiar with the standard method, but should not participate in the measurement process.

2.2.3 Laboratory

The laboratory personnel should be fully experienced with the test measurement method.

The laboratory shall undertake the analysis, adhering to the test procedures received from the executive officer. Any comments by the laboratory on the use of the test method should be reported to the executive officer. However, the procedures carried out by the laboratory should be those provided by the executive officer.

The laboratory shall comply with any requirements prescribed by the executive officer, including

— storage of the samples;

— date and order of carrying out the analysis.

The laboratory shall provide the executive officer with the results of analysis in a manner prescribed by the executive officer.

2.2.4 Statistician

The statistician shall receive from the executive officer the raw data obtained using the stated method and as reported in Table 1.

© ISO 2005 – All rights reserved

The statistician shall examine the data and apply any statistical test, preferentially the tests described in ISO 5725-2, to identify potential outliers. Statistical outliers shall be brought to the attention of the executive officer. The executive officer shall undertake an appropriate investigation to ascertain whether to retain, reject or modify any data.

The statistician shall carry out the statistical analyses, prepare graphical plots, and provide estimates of the means and variances (ISO 5725-2:1994, 7.1.2). The statistician shall summarize all the results of the statistical analyses in a report that shall be sent to the executive officer.

3 Critical examination of the data

3.1 Description of the data

3.1.1 Raw data

The data are presented as shown in Table 1. Tables 2 and 3 are derived from Table 1. Some validated statistical software packages may provide different presentations of the same information.

Table 1 — Collation of all raw data

Laboratory	Level						
	1	2	...	j	...	$q-1$	q
1							
2							
...							
i				y_{ij1} y_{ij2} y_{ijn}			
...							
p_{Lab}							

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e84...-ad42-16f4aeadc831/iso-tr-22971-2005

Table 2 — Collation of the mean values for each cell in Table 1

Laboratory	Level						
	1	2	...	j	...	$q-1$	q
1							
2							
...							
i				m_{ij}			
...							
p_{Lab}							

© ISO 2005 – All rights reserved

Table 3 — Collation of values indicating the spread [a] of the values for each cell in Table 1

Laboratory	Level						
	1	2	...	j	...	$q-1$	q
1							
2							
...							
i				s_{ij}			
...							
p_{Lab}							

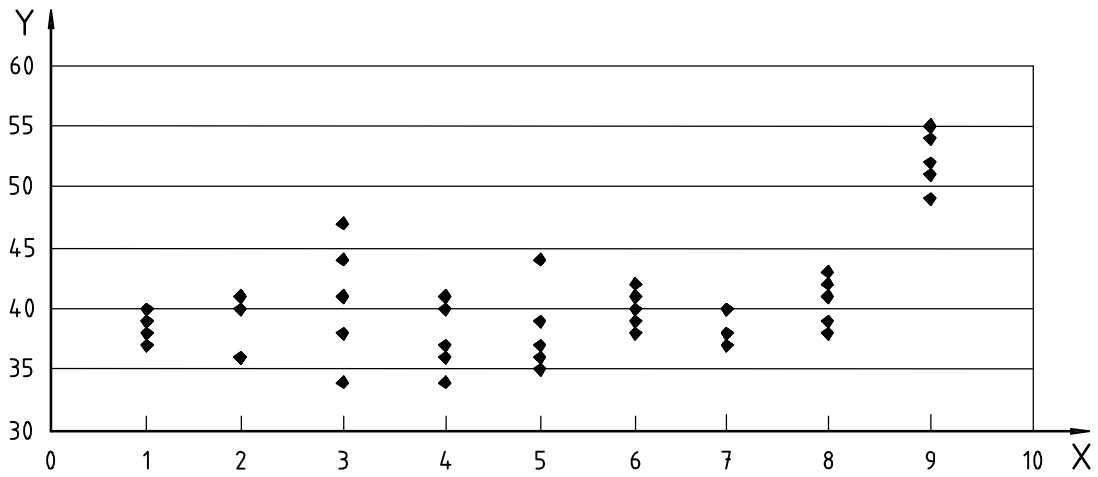
[a] The most common measure of spread is the standard deviation.

3.1.2 Graphical representation of the data

3.1.2.1 Results plotted versus laboratory number (raw data plot)

Before carrying out any tests to determine potential outliers, it is recommended that graphical plots of the raw data be made. In this way, an instant "picture" of the results can be depicted, for example, as shown in Figure 3 (which is based on ISO 5725-2:1994, Figures B.1 to B.4). A great deal of information can be obtained by a visual inspection of a graphical plot of the raw data, and an instant appraisal of the spread of data ascertained. Hence, an indication of the presence of outliers might be suggested, or unusual differences might become apparent, at particular levels of interest, simply by a visual inspection of the appropriate plot of data. For example, in Figure 3, the plot of results for laboratory 3 might suggest a larger-than-expected spread of results compared to all the other laboratories; hence, the overall repeatability will be affected. This possibility can be confirmed by Cochran's test. In addition, the results for laboratory 9 might suggest an outlier with respect to the mean value of the laboratory when compared to the other mean values for all the other laboratories. Hence, reproducibility might be affected and this can be confirmed by Grubbs' test.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-136e1-4267-ad42-16f4aeadc831/iso-tr-22971-2005



Key

X Laboratory identification

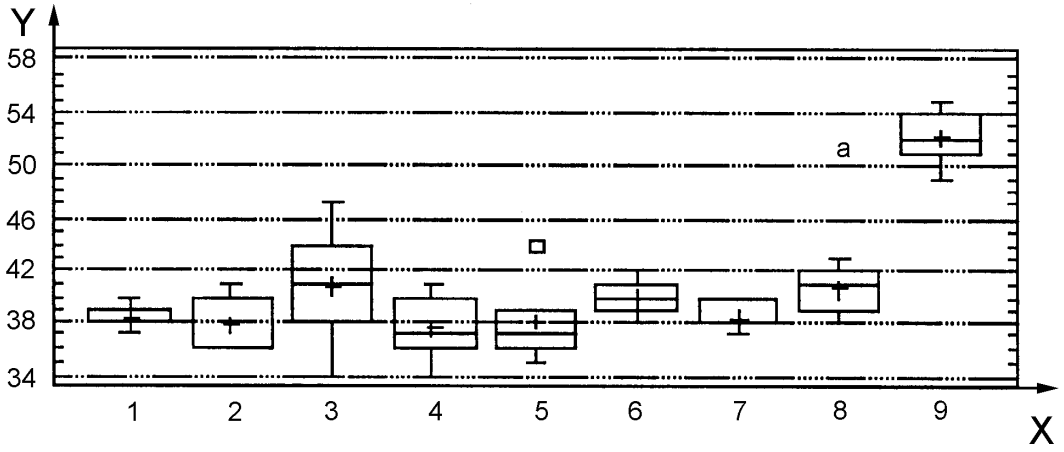
Y Results per laboratory

Figure 3 — Graphical representation of raw data for a particular level of interest

© ISO 2005 – All rights reserved

3.1.2.2 Boxplot ("box-and-whiskers" plot)

Where many results are reported, especially for a particular level of interest, a "box-and-whiskers" plot can reveal information similar to that in 3.1.2.1; for an example, see Figure 4. However, this type of plot, which is based on robust statistics including the determination of the median value, is not described in ISO 5725-2. It is, however, defined and illustrated in the examples in Clause 4, since these graphs are available in most statistical software packages.



Key

X Laboratory identification

Y Results per laboratory

a "+" indicates the average.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-10f4acadc831/iso-tr-22971-2005

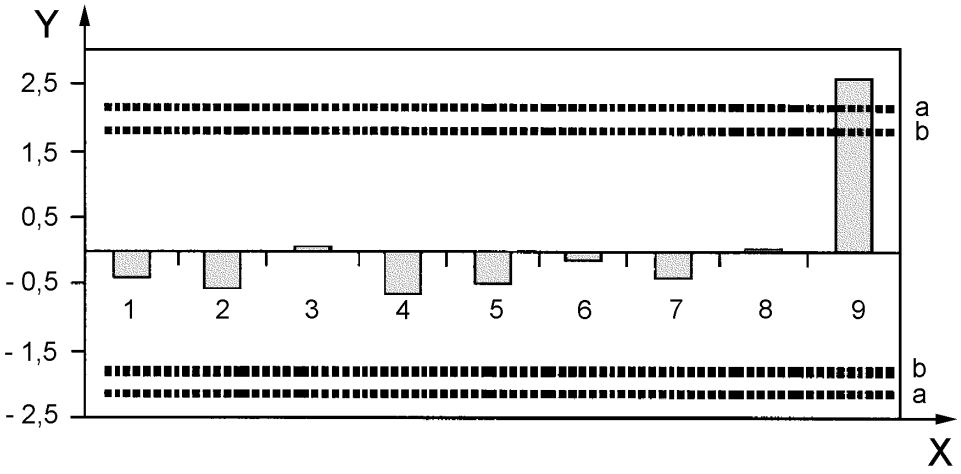
Figure 4 — "Box-and-whiskers" plot

3.1.2.3 Mandel's plots of h and k tests statistics

3.1.2.3.1 Mandel's h plot

For a particular level of interest, the mean values obtained for all the laboratories are used to calculate a single overall mean value. This value is then used to calculate Mandel's h statistic for all the laboratories for this level. This statistic is defined in ISO 5725-2:1994, Equation (6). This statistic is the ratio of the difference between the mean for a particular set of data and the mean of all sets of data, and the standard deviation of the means from all the sets of data. This quotient value is then plotted and compared with computed or tabulated ratio values obtained for 95 % and 99 % confidence levels. The same procedure is then used to calculate Mandel's h statistic for all the laboratories for all the other levels of interest (see Figure 5). It should be noted that both positive and negative values can be plotted.

© ISO 2005 – All rights reserved



Key

X Laboratory identification
Y Mendel's h

a 99 % confidence level
b 95 % confidence level

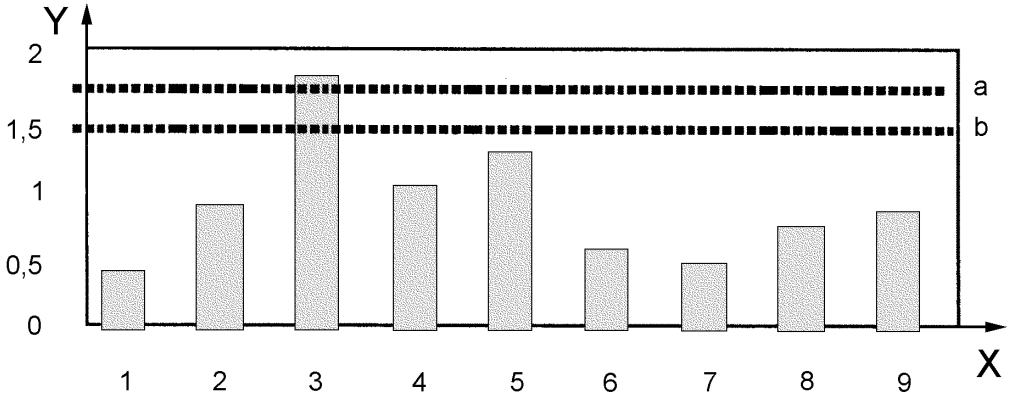
Figure 5 — Mendel's h plot

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.1.2.3.2 Mandel's k plot

For a particular level of interest, the standard deviations obtained for all the laboratories are used to calculate a mean standard deviation or pooled single standard deviation. This value is then used to calculate Mandel's k statistic for all the laboratories for this level. This statistic is defined in ISO 5725-2:1994, Equation (7). It is the quotient of the standard deviation of results and the mean or pooled standard deviation. This quotient value is then plotted and compared with computed or tabulated ratio values obtained for 95 % and 99 % confidence levels. The same procedure is then used to calculate Mandel's k statistic for all the laboratories for all the other levels of interest (see Figure 6). It should be noted that only positive values are plotted.

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23c1-4267-ad42-16f4aeeade831/iso-tr-22971-2005



Key

X Laboratory identification
Y Mendel's k

a 99 % confidence level
b 95 % confidence level

Figure 6 — Mendel's k plot

© ISO 2005 – All rights reserved

3.1.2.3.3 Graphical inspection

From the plots, individual results can be identified for each laboratory that might be considered different from the expected distribution of results. For example, the h plot for particular levels of interest for each laboratory might approach or exceed the computed Mandel's h statistic value at the 95 % or 99 % confidence level if the Grubbs' test shows outliers to be present. In addition, the k plot for particular levels of interest for each laboratory might approach or exceed the computed Mandel's k statistic value at the 95 % or 99 % confidence level if Cochran's test shows outliers to be present.

3.2 Tests for outliers

3.2.1 General points

3.2.1.1 Level of confidence

The treatment of outliers is dealt with in ISO 5725-2:1994, Clause 7, particularly 7.1 to 7.3. An outlier can be considered as a result which is sufficiently different from all other results to warrant further investigation. Depending on the type of distribution into which the results fit, a result that appears to be an outlier could, in reality, be a valid result. ISO 5725-2:1994, 7.3.2.1 and 7.3.3.2, recommends confidence levels of 95 % for outliers termed “stragglers”, and 99 % for outliers termed “statistical outliers”. For individual circumstances, the selection of 95 % and 99 % confidence levels means that one result in 20, and one result in 100, respectively, might be erroneously misinterpreted. Hence, this one result could occur by chance and the degree of confidence stated in ISO 5725-2 might not be appropriate for individual needs. This might represent a degree of acceptability that is not sufficient for certain purposes. This would mean that individual circumstances would merit individual consideration as to whether ISO 5725-2, in terms of the confidence levels used, should be applied.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.2.1.2 Basic assumptions

ISO/TR 22971:2005
https://standards.iteh.ai/catalog/standards/sist/e8406b2a-23e1-4267-ad42-16f4aeade831/iso-tr-22971-2005

In the tests used to determine the presence or absence of outliers, it is assumed that the results are distributed in a Gaussian manner (commonly referred to as a normal distribution; ISO 5725-2:1994, 1.4) or at least a single unimodal distribution (ISO 5725-2:1994, 7.3.1.7). Hence, before undertaking any test, especially one that involves a large number of results, a check to confirm this assumption should be made. It is also assumed (ISO 5725-2:1994, 1.3 and 5.1.1) that the number of results within each set of data (from each laboratory) is the same and that the number of results for each level of interest, or number of different samples, is the same. Thus, the results are “balanced”. If the results are not “balanced”, then it is recommended (ISO 5725-2:1994, 7.2.2) that results from appropriate sets of data be randomly discarded until a “balanced” situation is created. Although a “balanced” situation is preferred, it is recognized (even within the examples illustrated in ISO 5725-2) that “unbalanced” situations can be accommodated. It is further assumed (ISO 5725-1:1994, 4.4, and ISO 5725-2:1994, 7.3.3.3) that results are obtained under repeatability conditions. Hence, it can be assumed that the samples for a specific level of interest are homogeneous, identical in all respects, and analysed within a short period of time using the same reagents and calibration solutions. In theory, these criteria have to be satisfied before any tests can be used to establish the presence or absence of outliers.

3.2.1.3 Declaration of outliers

When carrying out the outlier tests, it should be understood that outliers should not be discarded or rejected purely from a statistical point of view. For each sample, the reason why the result is different from all the others should be investigated and identified. Outlier tests (based on the assumptions used) indicate whether there is sufficient statistical cause for an outlier; it will not indicate why it has occurred. It is only after thorough investigations have been undertaken to identify likely causes that data should be declared outliers and discarded.

When a particular level of interest has been analysed by Cochran's test, Grubbs' test or some other test and no outliers or further outliers are identified, then other levels of interest are similarly tested. If several outliers are identified in different levels of interest for a given set of data produced by a single laboratory, it may be necessary to consider whether all the data sets for all the levels of interest should be further investigated.

© ISO 2005 – All rights reserved