This document is a preview generated by EVS

STATISTILISED MEETODID LABORITEVAHELISTE VÕRDLUSTE TASEMEKATSETES KASUTAMISEKS

Statistical methods for use in proficiency testing by interlaboratory comparison (ISO 13528:2015)

EESTI STANDARDIKESKUS
ESTONIAN CENTRE FOR STANDARDISATION

EESTI STANDARDI EESSÕNA	NATIONAL FOREWORD
See Eesti standard EVS-ISO 13528:2017 „Statistilised meetodid laboritevaheliste võrdluste tasemekatsetes kasutamiseks“ sisaldab rahvusvahelise standardi ISO 13528:2015 „Statistical methods for use in proficiency testing by interlaboratory comparison“ identset ingliskeelset teksti.	This Estonian Standard EVS-ISO 13528:2017 consists of the identical English text of the International Standard ISO 13528:2015 „Statistical methods for use in proficiency testing by interlaboratory comparison“.
Ettepaneku rahvusvahelise standardi ümbertrüki meetodil ülevõtuks on esitanud EVS/TK 38, standardi avaldamist on korraldanud Eesti Standardikeskus.	Proposal to adopt the International Standard by reprint method has been presented by EVS/TK 38, the Estonian standard has been published by the Estonian Centre for Standardisation.
Standard EVS-ISO 13528:2017 on jõustunud sellekohase teate avaldamisega EVS Teataja 2017. aasta aprillikuu numbris.	Standard EVS-ISO 13528:2017 has been endorsed with a notification published in the April 2017 issue of the official bulletin of the Estonian Centre for Standardisation.
Standard on kättesaadav Eesti Standardikeskusest.	The standard is available from the Estonian Centre for Standardisation.

Käsitlusala

See rahvusvaheline standard esitab tasemekatsete korraldajatele statistiliste meetodite üksikasjalikud kirjeldused kujundamaks tasemekatsete skeeme ja analüüsimaks nendest katsetest saadud andmeid. See standard esitab soovitused saadud andmete tõlgendamiseks sellistes tasemekatsete skeemides osalejatele ja akrediteerimisasutustele.

Selles standardis esitatud protseduure saab rakendada, näitamaks et laborite, inspekteerimisasutuste ja isikute saadud mõõtetulemused on kooskõlas rahuldavale toimivusele esitatud kriteeriumitega.

Standard on kasutatav tasemekatsete korral, kus tulemusteks on nii kvantitatiivsed mõõtetulemused kui ka katseobjektide kvalitatiivsed vaatlustulemused.

MÄRKUS Selle standardi protseduurid võivad olla rakendatavad ekspertarvamuse hindamisel, kus arvamused või hinnangud esitatakse kujul, mida saab objektiivselt võrrelda sõltumatu tugiväärtuse või konsensusliku statistilise väärtusega. Näiteks kui klassifitseerida tasemekatse objekte inspekteerimise teel teadaolevatesse kategooriatesse või määratleda inspekteerimise teel, kas samast esialgsest allikast tekib katseobjekt või mitte, ja klassifitseerimise tulemusi võrrelda objektiivselt, võivad rakenduda selle standardi osad, mis seonduvad (kvalitatiivsete) vaikeomadustega.

This document is a preview generated by EVS

This document is a preview generated by EVS

Tagasisidet standardi sisu kohta on võimalik edastada, kasutades EVS-i veebilehel asuvat tagasiside vormi või saates e-kirja meiliaadressile standardiosakond@evs.ee.

ICS 03.120.30

Standardite reprodutseerimise ja levitamise õigus kuulub Eesti Standardikeskusele

Andmete paljundamine, taastekitamine, kopeerimine, salvestamine elektroonsesse süsteemi või edastamine ükskõik millises vormis või millisel teel ilma Eesti Standardikeskuse kirjaliku loata on keelatud.

Kui Teil on küsimusi standardite autorikaitse kohta, võtke palun ühendust Eesti Standardikeskusega:
Koduleht www.evs.ee; telefon 605 5050; e-post info@evs.ee

The right to reproduce and distribute standards belongs to the Estonian Centre for Standardisation

No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, without a written permission from the Estonian Centre for Standardisation.

If you have any questions about copyright, please contact Estonian Centre for Standardisation:
Homepage www.evs.ee; phone +372 605 5050; e-mail info@evs.ee

Contents

Page

Foreword ... **v**
0 Introduction ... **vii**
1 Scope ... **1**
2 Normative references ... **1**
3 Terms and definitions ... **1**
4 General principles ... **4**
4.1 General requirements for statistical methods ... 4
4.2 Basic model ... 5
4.3 General approaches for the evaluation of performance ... 5
5 Guidelines for the statistical design of proficiency testing schemes ... **6**
5.1 Introduction to the statistical design of proficiency testing schemes ... 6
5.2 Basis of a statistical design ... 6
5.3 Considerations for the statistical distribution of results ... 7
5.4 Considerations for small numbers of participants ... 8
5.5 Guidelines for choosing the reporting format ... 8
6 Guidelines for the initial review of proficiency testing items and results ... **10**
6.1 Homogeneity and stability of proficiency test items ... 10
6.2 Considerations for different measurement methods ... 11
6.3 Blunder removal ... 11
6.4 Visual review of data ... 11
6.5 Robust statistical methods ... 12
6.6 Outlier techniques for individual results ... 12
7 Determination of the assigned value and its standard uncertainty ... **13**
7.1 Choice of method of determining the assigned value ... 13
7.2 Determining the uncertainty of the assigned value ... 14
7.3 Formulation ... 15
7.4 Certified reference material ... 15
7.5 Results from one laboratory ... 16
7.6 Consensus value from expert laboratories ... 17
7.7 Consensus value from participant results ... 18
7.8 Comparison of the assigned value with an independent reference value ... 19
8 Determination of criteria for evaluation of performance ... **20**
8.1 Approaches for determining evaluation criteria ... 20
8.2 By perception of experts ... 20
8.3 By experience from previous rounds of a proficiency testing scheme ... 20
8.4 By use of a general model ... 21
8.5 Using the repeatability and reproducibility standard deviations from a previous collaborative study of precision of a measurement method ... 22
8.6 From data obtained in the same round of a proficiency testing scheme ... 22
8.7 Monitoring interlaboratory agreement ... 23
9 Calculation of performance statistics ... **23**
9.1 General considerations for determining performance ... 23
9.2 Limiting the uncertainty of the assigned value ... 24
9.3 Estimates of deviation (measurement error) ... 25
9.4 z scores ... 26
9.5 z' scores ... 27
9.6 Zeta scores (ζ) ... 28
9.7 E_n scores ... 29
9.8 Evaluation of participant uncertainties in testing ... 29
9.9 Combined performance scores ... 30

This document is a preview generated by EVS

© ISO 2015 – All rights reserved

10 Graphical methods for describing performance scores 31
10.1 Application of graphical methods 31
10.2 Histograms of results or performance scores 31
10.3 Kernel density plots 32
10.4 Bar-plots of standardized performance scores 33
10.5 Youden Plot 33
10.6 Plots of repeatability standard deviations 34
10.7 Split samples 35
10.8 Graphical methods for combining performance scores over several rounds of a proficiency testing scheme 36

11 Design and analysis of qualitative proficiency testing schemes (including nominal and ordinal properties) 37
11.1 Types of qualitative data 37
11.2 Statistical design 37
11.3 Assigned values for qualitative proficiency testing schemes 38
11.4 Performance evaluation and scoring for qualitative proficiency testing schemes 39

Annex A (normative) **Symbols 41**

Annex B (normative) **Homogeneity and stability of proficiency test items 43**

Annex C (normative) **Robust analysis 51**

Annex D (informative) **Additional guidance on statistical procedures 63**

Annex E (informative) **Illustrative examples 67**

Bibliography 88

This document is a preview generated by EVS

 © ISO 2015 – 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.

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 on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

This second edition of ISO 13528 cancels and replaces the first edition (ISO 13528:2005), of which it constitutes a technical revision. This second edition provides changes to bring the document into harmony with ISO/IEC 17043:2010, which replaced ISO Guide 43-1:1997. It follows a revised structure, to describe better the process of the design, analysis, and reporting of proficiency testing schemes. It also eliminates some procedures that are no longer considered to be appropriate, and adds or revises some other sections to be consistent with ISO/IEC 17043 and to provide clarity and correct minor errors. New sections have been added for qualitative data and additional robust statistical methods.

This corrected version of ISO 13528:2015 incorporates the following corrections:

— 7.5.2.2, equation (5):"U_{char}" has been replaced with "u_{char}";

— 9.8.2, NOTE: reference to "E.3" has been replaced with a reference to "E.4";

— 10.3.2, iv, equation (19): an "addition" sign has been added between "q_{min}" and "(i-1)";

— B.2.3, b), line before Table B.1: Reference "[33]" has been replaced with Reference "[32]";

— B.2.3, Table B.1, first row, first column: "gm" has been replaced with "g";

— B.2.3, paragraph below Table B.1: in the first formula,"F_{2m}" has been replaced with "F_m" and subscript for "-1" has been removed; in the second formula, subscript for "-1" has been removed;

— B.2.3 b), NOTE: the equation for F_1 is now divided by "(g-1)";

— B.3, equations (B.7), (B.14) and (B.16): the terms s_x^2 and s_x have been replaced with $s_{\bar{x}}^2$ and $s_{\bar{x}}$; in addition, in equation (B.16) the square root symbol has been moved outside the "max (0, ...)" bracket;

© ISO 2015 – All rights reserved

This document is a preview generated by EVS

— B.3, equation (B.8): "s_t^2" has been replaced with "w_t^2";

— C.3.1, NOTE 2, first line: "the" has been removed before "identical";

— C.3.1, paragraph after equation (C.10), second line: the words "the modified data in" have been deleted.

— C.3.1, last note: "NOTE" has been replaced with "NOTE 3", and Reference to "E.3 and E.4" has been replaced with a Reference to "E.1 and E.3";

The following minor editorial corrections have been implemented for consistency throughout the document:

— 8.3.1, third bullet, last line: the first occurrence of "approved" (after "more") has been deleted;

— 8.6.1, first line: "$,\sigma_{pt},$" has been replaced with "(σ_{pt})" (for presentation consistency);

— B.4.1.2, second bullet, second line: the word "samples" has been replaced with "proficiency testing items" (for terminological consistency)

— Annexes D and E, titles: the first letters in all words after the first one is now in lower case (for presentation consistency).

This document is a preview generated by EVS

 © ISO 2015 – All rights reserved

0 Introduction

0.1 The purposes of proficiency testing

Proficiency testing involves the use of interlaboratory comparisons to determine the performance of participants (which may be laboratories, inspection bodies, or individuals) for specific tests or measurements, and to monitor their continuing performance. There are a number of typical purposes of proficiency testing, as described in the Introduction to ISO/IEC 17043:2010. These include the evaluation of laboratory performance, the identification of problems in laboratories, establishing effectiveness and comparability of test or measurement methods, the provision of additional confidence to laboratory customers, validation of uncertainty claims, and the education of participating laboratories. The statistical design and analytical techniques applied must be appropriate for the stated purpose(s).

0.2 Rationale for scoring in proficiency testing schemes

A variety of scoring strategies is available and in use for proficiency testing. Although the detailed calculations differ, most proficiency testing schemes compare the participant's deviation from an assigned value with a numerical criterion which is used to decide whether or not the deviation represents cause for concern. The strategies used for value assignment and for choosing a criterion for assessment of the participant deviations are therefore critical. In particular, it is important to consider whether the assigned value and criterion for assessing deviations should be independent of participant results, or should be derived from the results submitted. In this Standard, both strategies are provided for. However, attention is drawn to the discussion that will be found in sections 7 and 8 of the advantages and disadvantages of choosing assigned values or criteria for assessing deviations that are not derived from the participant results. It will be seen that in general, choosing assigned values and assessment criteria independently of participant results offers advantages. This is particularly the case for the criterion used to assess deviations from the assigned value – such as the standard deviation for proficiency assessment or an allowance for measurement error – for which a consistent choice based on suitability for a particular end use of the measurement results, is especially useful.

0.3 ISO 13528 and ISO/IEC 17043

ISO 13528 provides support for the implementation of ISO/IEC 17043 particularly, on the requirements for the statistical design, validation of proficiency test items, review of results, and reporting summary statistics. Annex B of ISO/IEC 17043:2010 briefly describes the general statistical methods that are used in proficiency testing schemes. This International Standard is intended to be complementary to ISO/IEC 17043, providing detailed guidance that is lacking in that document on particular statistical methods for proficiency testing.

The definition of proficiency testing in ISO/IEC 17043 is repeated in ISO 13528, with the Notes that describe different types of proficiency testing and the range of designs that can be used. This Standard cannot specifically cover all purposes, designs, matrices and measurands. The techniques presented in ISO 13528 are intended to be broadly applicable, especially for newly established proficiency testing schemes. It is expected that statistical techniques used for a particular proficiency testing scheme will evolve as the scheme matures; and the scores, evaluation criteria, and graphical techniques will be refined to better serve the specific needs of a target group of participants, accreditation bodies, and regulatory authorities.

ISO 13528 incorporates published guidance for the proficiency testing of chemical analytical laboratories [32] but additionally includes a wider range of procedures to permit use with valid measurement methods and qualitative identifications. This revision of ISO 13528:2005 contains most of the statistical methods and guidance from the first edition, extended as necessary by the previously referenced documents and the extended scope of ISO/IEC 17043. ISO/IEC 17043 includes proficiency testing for individuals and inspection bodies, and Annex B, which includes considerations for qualitative results.

This Standard includes statistical techniques that are consistent with other International Standards, particularly those of TC69 SC6, notably the ISO 5725 series of standards on *Accuracy: trueness and*

This document is a preview generated by EVS

© ISO 2015 – All rights reserved

precision. The techniques are also intended to reflect other international standards, where appropriate, and are intended to be consistent with ISO/IEC Guide 98-3 (GUM) and ISO/IEC Guide 99 (VIM).

0.4 Statistical expertise

ISO/IEC 17043:2010 requires that in order to be competent, a proficiency testing provider shall have access to statistical expertise and shall authorize specific personnel to conduct statistical analysis. Neither ISO/IEC 17043 nor this International Standard can specify further what that necessary expertise is. For some applications an advanced degree in statistics is useful, but usually the needs for expertise can be met by individuals with technical expertise in other areas, who are familiar with basic statistical concepts and have experience or training in the common techniques applicable to the analysis of data from proficiency testing schemes. If an individual is charged with statistical design and/or analysis, it is very important that this person has experience with interlaboratory comparisons, even if that person has an advanced degree in statistics. Conventional advanced statistical training often does not include exercises with interlaboratory comparisons, and the unique causes of measurement error that occur in proficiency testing can seem obscure. The guidance in this International Standard cannot provide all the necessary expertise to consider all applications, and cannot replace the experience gained by working with interlaboratory comparisons.

0.5 Computer software

Computer software that is needed for statistical analysis of proficiency testing data can vary greatly, ranging from simple spread sheet arithmetic for small proficiency testing schemes using known reference values to sophisticated statistical software used for statistical methods reliant on iterative calculations or other advanced numerical methods. Most of the techniques in this International Standard can be accomplished by conventional spread sheet applications, perhaps with customised routines for a particular scheme or analysis; some techniques will require computer applications that are freely available (at the time of publication of this Standard). In all cases, the users should verify the accuracy of their calculations, especially when special routines have been entered by the user. However, even when the techniques in this International Standard are appropriate and correctly implemented by adequate computer applications, they cannot be applied without attention from an individual with technical and statistical expertise that is sufficient to identify and investigate anomalies that can occur in any round of proficiency testing.

This document is a preview generated by EVS

© ISO 2015 – All rights reserved

Statistical methods for use in proficiency testing by interlaboratory comparison

This document is a preview generated by EVS

1 Scope

This International Standard provides detailed descriptions of statistical methods for proficiency testing providers to use to design proficiency testing schemes and to analyse the data obtained from those schemes. This Standard provides recommendations on the interpretation of proficiency testing data by participants in such schemes and by accreditation bodies.

The procedures in this Standard can be applied to demonstrate that the measurement results obtained by laboratories, inspection bodies, and individuals meet specified criteria for acceptable performance.

This Standard is applicable to proficiency testing where the results reported are either quantitative measurements or qualitative observations on test items.

NOTE The procedures in this Standard may also be applicable to the assessment of expert opinion where the opinions or judgments are reported in a form which may be compared objectively with an independent reference value or a consensus statistic. For example, when classifying proficiency test items into known categories by inspection - or in determining by inspection whether proficiency test items arise, or do not arise, from the same original source - and the classification results are compared objectively, the provisions of this Standard that relate to nominal (qualitative) properties may apply.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO Guide 30, *Reference materials — Selected terms and definitions*

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*

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

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

ISO/IEC 17043, *Conformity assessment — General requirements for proficiency testing*

ISO/IEC Guide 99, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM)*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 3534-1, ISO 3534-2, ISO 5725-1, ISO/IEC 17043, ISO/IEC Guide 99, ISO Guide 30, and the following apply. In the case of differences between these references on the use of terms, definitions in ISO 3534 parts 1-2 apply. Mathematical symbols are listed in Annex A.

3.1
interlaboratory comparison
organization, performance and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions

© ISO 2015 – All rights reserved