This document is a preview generated by EVS

STATISTILISED MEETODID LABORITEVAHELISE VÕRDLUSE TEEL TEOSTATAVATEL TASEMEKATSETEL KASUTAMISEKS

Statistical methods for use in proficiency testing by interlaboratory comparison (ISO 13528:2022, identical)

EESTI STANDARDI EESSÕNA	NATIONAL FOREWORD
See Eesti standard EVS-ISO 13528:2023 sisaldab rahvusvahelise standardi ISO 13528:2022 „Statistical methods for use in proficiency testing by interlaboratory comparison“ identset ingliskeelset teksti. Ettepaneku rahvusvahelise standardi ümbertrüki meetodil ülevõtuks on esitanud EVS/TK 38, standardi avaldamist on korraldanud Eesti Standardimis- ja Akrediteerimiskeskus. Standard EVS-ISO 13528:2023 on jõustunud sellekohase teate avaldamisega EVS Teatajas. Standard on kättesaadav Eesti Standardimis- ja Akrediteerimiskeskusest.	This Estonian Standard EVS-ISO 13528:2023 consists of the identical English text of the International Standard ISO 13528:2022 „Statistical methods for use in proficiency testing by interlaboratory comparison“. 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 and Accreditation. Standard EVS-ISO 13528:2023 has been endorsed with a notification published in the official bulletin of the Estonian Centre for Standardisation and Accreditation. This standard is available from the Estonian Centre for Standardisation and Accreditation.

Käsitlusala

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

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

Dokument on kasutatav tasemekatsete korral, kus tulemused on kas kvantitatiivsed mõõtetulemused või katseobjektide kvalitatiivsed vaatlustulemused.

MÄRKUS Selle dokumendi protseduure saab rakendada eksperdiarvamuse 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, saab rakendada selle dokumendi osi, 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 Standardimis- ja Akrediteerimiskeskusele

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

Kui Teil on küsimusi standardite autoriõiguse kaitse kohta, võtke palun ühendust Eesti Standardimis- ja Akrediteerimiskeskusega: 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 and Accreditation

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 and Accreditation.

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

This document is a preview generated by EVS

[Blank page]

Contents

Page

Foreword v
0 Introduction vi
1 Scope 1
2 Normative references 1
3 Terms and definitions 1
4 General principles 5
4.1 General requirements for statistical methods 5
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 9
5.5.1 General requirements for reporting format 9
5.5.2 Reporting of replicate measurements 9
5.5.3 Reporting of 'less than' or 'greater than' a limit (censored data) 9
5.5.4 Number of significant digits 10
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 12
6.3 Blunder removal 12
6.4 Visual review of data 12
6.5 Robust statistical methods 13
6.6 Outlier techniques for individual results 13
7 Determination of the assigned value and its standard uncertainty 14
7.1 Choice of method of determining the assigned value 14
7.2 Determining the uncertainty of the assigned value 15
7.3 Formulation 16
7.4 Certified reference material 16
7.5 Results from one laboratory 17
7.6 Consensus value from expert laboratories 18
7.7 Consensus value from participant results 19
7.8 Comparison of the assigned value with an independent reference value 20
8 Determination of criteria for evaluation of performance 21
8.1 Approaches for determining evaluation criteria 21
8.2 By perception of experts 21
8.3 By experience from previous rounds of a proficiency testing scheme 21
8.4 By use of a general model 22
8.5 Using the repeatability and reproducibility standard deviations from a previous collaborative study of precision of a measurement method 23
8.6 From data obtained in the same round of a proficiency testing scheme 23
8.7 Monitoring interlaboratory agreement 24
9 Calculation of performance statistics 25
9.1 General considerations for determining performance 25
9.2 Limiting the uncertainty of the assigned value 25

This document is a preview generated by EVS

© ISO 2022 – All rights reserved

9.3 Estimates of deviation (measurement error) 26
9.4 z scores 27
9.5 z' scores 28
9.6 Zeta scores (ζ) 29
9.7 E_n scores 30
9.8 Evaluation of participant uncertainties in testing 31
9.9 Combined performance scores 32

10 Graphical methods for describing performance scores 33
10.1 Application of graphical methods 33
10.2 Histograms of results or performance scores 33
10.3 Kernel density plots 34
10.4 Bar-plots of standardized performance scores 35
10.5 Youden plot 35
10.6 Plots of repeatability standard deviations 36
10.7 Split samples 37
10.8 Graphical methods for combining performance scores over several rounds of a proficiency testing scheme 38

11 Design and analysis of qualitative proficiency testing schemes (including nominal and ordinal properties) 39
11.1 Types of qualitative data 39
11.2 Statistical design 39
11.3 Assigned values for qualitative proficiency testing schemes 40
11.4 Performance evaluation and scoring for qualitative proficiency testing schemes 41

Annex A (normative) **Symbols 44**

Annex B (informative) **Homogeneity and stability of proficiency test items 46**

Annex C (informative) **Robust analysis 54**

Annex D (informative) **Additional guidance on statistical procedures 65**

Annex E (informative) **Illustrative examples 70**

Annex F (informative) **Example of computer code for plotting and resampling analysis ("bootstrapping") of PT results 95**

Bibliography 96

This document is a preview generated by EVS

 © ISO 2022 – 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 third edition of ISO 13528 cancels and replaces the second edition (ISO 13528:2015), of which it constitutes an minor revision. The changes are as follows:

— notes have been added to 10.1, 10.4.3 and 10.5.3 to draw attention to additional graphical techniques that can assist in meeting the provisions of 10.1;

— Formulae B.4 and B.8 have been corrected to use s_t^2 instead of w_t^2;

— Formula B.16 has been corrected so that the term inside the square root is always non-negative;

— in Table C.2, the correction factor associated with $p = 2$ has been corrected to read 0,3994;

— additional literature references to the source of values in Table C.2 have been added to the Bibliography and referenced from Notes 1 and 2 of C.5.2.1;

— font styles (Italic or Roman) have been amended throughout for consistency in formulae.

This document is a preview generated by EVS

© ISO 2022 – All rights reserved

0Introduction

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. 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 shall 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 document, both strategies are provided for. However, attention is drawn to the discussion that will be found in Clauses 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

This document 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. ISO/IEC 17043:2010, Annex B, briefly describes the general statistical methods that are used in proficiency testing schemes. This document 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 this document, with the notes that describe different types of proficiency testing and the range of designs that can be used. This document cannot specifically cover all purposes, designs, matrices and measurands. The techniques presented in this document 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.

This document 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. The revision of this document 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, including ISO/IEC 17043:2010, Annex B, which includes considerations for qualitative results.

This document is a preview generated by EVS

 © ISO 2022 – All rights reserved

This document 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 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 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 document 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 responsible for 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 document 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 document can be accomplished by conventional spread sheet applications, perhaps with customised routines for a particular proficiency testing scheme or analysis; some techniques will require computer applications that are freely available. In all cases, the users are expected to verify the validity and accuracy of their calculations, especially when special routines have been entered by the user. However, even when the techniques in this document 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 understand the nature of the applications and the statistical assumptions, and to identify and investigate anomalies that can occur in any round of a proficiency testing scheme.

This document is a preview generated by EVS

© ISO 2022 – All rights reserved

This document is a preview generated by EVS

[Blank page]

Statistical methods for use in proficiency testing by interlaboratory comparisob

This document is a preview generated by EVS

1 Scope

This document 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 document provides recommendations on the interpretation of proficiency testing data by participants in such proficiency testing schemes and by accreditation bodies.

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

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

NOTE The procedures in this document can also be applied for the assessment of expert opinion where the opinions or judgments are reported in a form which can 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 document that relate to nominal (qualitative) properties can be applied.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: 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 Guide 30, *Reference materials — Selected terms and definitions*

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 3534-1 ISO 3534-2 apply. Mathematical symbols are listed in Annex A.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

© ISO 2022 – All rights reserved