ISO 13528:2015-08 (E)

Statistical methods for use in proficiency testing by interlaboratory comparison

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 () 28
9.7 En scores 29
9.8 Evaluation of participant uncertainties in testing 29
9.9 Combined performance scores 30

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