

DIN ISO 13528:2020-09 (D/E)

Statistische Verfahren für Eignungsprüfungen durch Ringversuche (ISO 13528:2015, korrigierte Fassung 2016-10-15); Text Deutsch und Englisch

Statistical methods for use in proficiency testing by interlaboratory comparison (ISO 13528:2015, Corrected version 2016-10-15); Text in German and English

| Inhalt | Seite |
|------------------------------------------------------------------------------------------------------------------------------|--------------|
| Nationales Vorwort | 4 |
| Nationaler Anhang NA (informativ) Literaturhinweise | 5 |
| Vorwort | 6 |
| 0 Einleitung | 8 |
| 0.1 Zweck von Eignungsprüfungen..... | 8 |
| 0.2 Begründung für die Bewertung durch Scores in Eignungsprüfungsprogrammen..... | 8 |
| 0.3 ISO 13528 und ISO/IEC 17043 | 8 |
| 0.4 Statistischer Sachverstand | 9 |
| 0.5 Computersoftware | 9 |
| 1 Anwendungsbereich | 10 |
| 2 Normative Verweisungen | 10 |
| 3 Begriffe | 10 |
| 4 Allgemeine Grundsätze | 14 |
| 4.1 Allgemeine Anforderungen an statistische Verfahren | 14 |
| 4.2 Grundmodell | 14 |
| 4.3 Allgemeine Ansätze für die Leistungsbewertung | 15 |
| 5 Leitlinien für die statistische Planung von Eignungsprüfungsprogrammen | 15 |
| 5.1 Einleitung zur statistischen Planung von Eignungsprüfungsprogrammen..... | 15 |
| 5.2 Grundlage einer statistischen Planung | 15 |
| 5.3 Überlegungen zur statistischen Verteilung der Ergebnisse | 16 |
| 5.4 Überlegungen für kleine Teilnehmeranzahlen | 17 |
| 5.5 Leitlinien zur Auswahl des Berichtsformats | 18 |
| 6 Leitlinien für die anfängliche Überprüfung von Prüfgegenständen für die Eignungsprüfungen und von Ergebnissen | 20 |
| 6.1 Homogenität und Stabilität von Prüfgegenständen für die Eignungsprüfung | 20 |
| 6.2 Betrachtungen für verschiedene Messverfahren..... | 21 |
| 6.3 Entfernung grober Fehler | 22 |
| 6.4 Visuelle Überprüfung von Daten | 22 |
| 6.5 Robuste statistische Verfahren..... | 22 |
| 6.6 Ausreißertechniken für Einzelergebnisse | 23 |
| 7 Bestimmung des zugewiesenen Werts und seiner Standardunsicherheit | 24 |
| 7.1 Auswahl des Verfahrens zur Bestimmung des zugewiesenen Werts | 24 |
| 7.2 Bestimmung der Unsicherheit des zugewiesenen Werts | 24 |
| 7.3 Bekannte Werte aus der Probenherstellung..... | 26 |
| 7.4 Zertifiziertes Referenzmaterial | 27 |
| 7.5 Ergebnisse aus einem einzelnen Laboratorium | 27 |
| 7.6 Konsenswert von Expertenlaboratorien | 28 |

| | | |
|------|---------------------------------------------------------------------------------------------------------------------------------|-----|
| 7.7 | Konsenswert aus Teilnehmerergebnissen | 29 |
| 7.8 | Vergleich des zugewiesenen Werts mit einem unabhängigen Referenzwert..... | 31 |
| 8 | Bestimmung von Kriterien für die Leistungsbewertung..... | 32 |
| 8.1 | Ansätze für die Bestimmung von Bewertungskriterien | 32 |
| 8.2 | Werte aus Erkenntnissen von Experten | 32 |
| 8.3 | Werte aus Erfahrungen aus früheren Runden eines Eignungsprüfungsprogramms..... | 33 |
| 8.4 | Wert aus einem allgemeinen Modell | 33 |
| 8.5 | Werte aus Wiederhol- und Vergleichstandardabweichungen eines früheren Ringversuchs zur Präzision eines Messverfahrens..... | 34 |
| 8.6 | Werte aus Daten, die in derselben Runde eines Eignungsprüfungsprogramms erzielt wurden..... | 35 |
| 8.7 | Überwachung der Übereinstimmung zwischen Laboratorien | 36 |
| 9 | Berechnung von Leistungskenngrößen | 37 |
| 9.1 | Allgemeine Erwägungen für die Leistungsbestimmung | 37 |
| 9.2 | Begrenzung der Unsicherheit des zugewiesenen Werts | 37 |
| 9.3 | Schätzwerte der Abweichung (Messabweichung) | 38 |
| 9.4 | z-Scores..... | 39 |
| 9.5 | z'-Scores | 40 |
| 9.6 | Zeta-Scores (ζ) | 41 |
| 9.7 | E_n -Scores..... | 43 |
| 9.8 | Bewertung der Teilnehmerunsicherheiten bei der Prüfung | 43 |
| 9.9 | Kombinierte Leistungs-Scores | 45 |
| 10 | Grafische Verfahren zur Beschreibung von Leistungswerten | 45 |
| 10.1 | Anwendung grafischer Verfahren..... | 45 |
| 10.2 | Histogramme von Ergebnissen oder Leistungswerten | 46 |
| 10.3 | Kerndichtediagramme | 46 |
| 10.4 | Säulendiagramme von standardisierten Leistungswerten | 48 |
| 10.5 | Youden-Diagramm..... | 48 |
| 10.6 | Diagramme von Wiederholstandardabweichungen..... | 49 |
| 10.7 | Geteilte Proben | 50 |
| 10.8 | Grafische Verfahren zur Kombination von Leistungswerten über mehrere Runden eines Eignungsprüfungsprogramms | 51 |
| 11 | Planung und Analyse von qualitativen Eignungsprüfungsprogrammen (einschließlich nominaler und ordinaler Eigenschaften) | 52 |
| 11.1 | Arten von qualitativen Daten | 52 |
| 11.2 | Statistische Planung | 52 |
| 11.3 | Zugewiesene Werte für qualitative Eignungsprüfungsprogramme | 53 |
| 11.4 | Leistungsbewertung und Bewertung durch Scores bei qualitativen Eignungsprüfungsprogrammen | 55 |
| | Anhang A (normativ) Symbole und Abkürzungen | 57 |
| | Anhang B (normativ) Homogenität und Stabilität von Prüfgegenständen für die Eignungsprüfung | 59 |
| | Anhang C (normativ) Robuste Analyse | 68 |
| | Anhang D (informativ) Zusätzlicher Leitfaden zu statistischen Verfahren | 80 |
| | Anhang E (informativ) Erläuternde Beispiele | 85 |
| | Literaturhinweise | 110 |

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 <i>z</i> scores | 26 |
| 9.5 <i>z'</i> 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 |

| | | |
|---------------------|------------------------------------------------------------------------------------------------------------------|-----------|
| 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 |