

# E DIN ISO 13528:2024-04 (D/E)

Erscheinungsdatum: 2024-03-15

Statistische Verfahren für Eignungsprüfungen durch Vergleiche zwischen  
Laboratorien (ISO 13528:2022); Text Deutsch und Englisch

Statistical methods for use in proficiency testing by interlaboratory comparison (ISO  
13528:2022); Text in German and English

---

Inhalt	Seite
Nationales Vorwort .....	7
Nationaler Anhang NA (informativ) Literaturhinweise .....	9
Vorwort .....	10
0 Einleitung.....	11
0.1 Zweck von Eignungsprüfungen.....	11
0.2 Begründung für die Bewertung durch Scores in Eignungsprüfungsprogrammen.....	11
0.3 ISO 13528 und ISO/IEC 17043 .....	11
0.4 Statistischer Sachverstand .....	12
0.5 Computersoftware .....	12
1 Anwendungsbereich.....	13
2 Normative Verweisungen .....	13
3 Begriffe .....	13
4 Allgemeine Grundsätze .....	17
4.1 Allgemeine Anforderungen an statistische Verfahren.....	17
4.2 Grundmodell.....	17
4.3 Allgemeine Ansätze für die Leistungsbewertung .....	18
5 Leitlinien für das statistische Modell von Eignungsprüfungsprogrammen .....	18
5.1 Einleitung zum statistischen Modell von Eignungsprüfungsprogrammen.....	18
5.2 Grundlage eines statistischen Modells.....	19
5.3 Überlegungen zur statistischen Verteilung der Ergebnisse .....	20
5.4 Überlegungen für kleine Teilnehmeranzahlen .....	21
5.5 Leitlinien zur Auswahl des Berichtsformats .....	21
5.5.1 Allgemeine Anforderungen an das Berichtsformat.....	21
5.5.2 Berichte über Wiederholmessungen .....	22
5.5.3 Berichte von „kleiner als“ oder „größer als“ ein Grenzwert (zensierte Daten ) .....	22
5.5.4 Anzahl der signifikanten Stellen .....	23
6 Leitlinien für die anfängliche Überprüfung von Eignungsprüfungsgegenständen und von Ergebnissen.....	23
6.1 Homogenität und Stabilität von Eignungsprüfungsgegenständen.....	23
6.2 Betrachtungen für verschiedene Messverfahren.....	25
6.3 Entfernung grober Fehler .....	25
6.4 Visuelle Überprüfung von Daten .....	25
6.5 Robuste statistische Verfahren.....	26
6.6 Ausreißertechniken für Einzelergebnisse .....	27
7 Bestimmung des zugewiesenen Werts und seiner Standardunsicherheit .....	28
7.1 Auswahl des Verfahrens zur Bestimmung des zugewiesenen Werts .....	28
7.2 Bestimmung der Unsicherheit des zugewiesenen Werts.....	28
7.3 Formulierung.....	29
7.4 Zertifiziertes Referenzmaterial .....	30

7.5	Ergebnisse aus einem einzelnen Laboratorium.....	30
7.6	Konsenswert von Expertenlaboratorien.....	32
7.7	Konsenswert aus Teilnehmerergebnissen.....	33
7.8	Vergleich des zugewiesenen Werts mit einem unabhängigen Referenzwert.....	34
8	Bestimmung von Kriterien für die Leistungsbewertung.....	36
8.1	Ansätze für die Bestimmung von Bewertungskriterien.....	36
8.2	Werte aus Erkenntnissen von Experten.....	36
8.3	Werte aus Erfahrungen aus früheren Runden eines Eignungsprüfungsprogramms.....	36
8.4	Wert aus einem allgemeinen Modell.....	37
8.5	Werte aus Wiederhol- und Vergleichstandardabweichungen eines früheren Vergleichs zwischen Laboratorien zur Präzision eines Messverfahrens.....	38
8.6	Werte aus Daten, die in derselben Runde eines Eignungsprüfungsprogramms erzielt wurden.....	38
8.7	Überwachung der Übereinstimmung zwischen Laboratorien.....	40
9	Berechnung von Leistungskenngrößen.....	40
9.1	Allgemeine Erwägungen für die Leistungsbestimmung.....	40
9.2	Begrenzung der Unsicherheit des zugewiesenen Werts.....	40
9.3	Schätzwerte der Abweichung (Messabweichung).....	41
9.4	z-Scores.....	43
9.5	z'-Scores.....	44
9.6	Zeta-Scores ( $\zeta$ ).....	45
9.7	$E_n$ -Scores.....	46
9.8	Bewertung der Teilnehmerunsicherheiten bei der Prüfung.....	47
9.9	Kombinierte Leistungs-Scores.....	48
10	Grafische Verfahren zur Beschreibung von Leistungs-Scores.....	49
10.1	Anwendung grafischer Verfahren.....	49
10.2	Histogramme von Ergebnissen oder Leistungs-Scores.....	49
10.3	Kerndichtediagramme.....	50
10.4	Säulendiagramme von standardisierten Leistungs-Scores.....	52
10.5	Youden-Diagramm.....	52
10.6	Diagramme von Wiederholstandardabweichungen.....	53
10.7	Geteilte Proben.....	54
10.8	Grafische Verfahren zur Kombination von Leistungs-Scores über mehrere Runden eines Eignungsprüfungsprogramms.....	55
11	Planung und Analyse von qualitativen Eignungsprüfungsprogrammen (einschließlich nominaler und ordinaler Eigenschaften).....	56
11.1	Arten von qualitativen Daten.....	56
11.2	Statistisches Modell.....	57
11.3	Zugewiesene Werte für qualitative Eignungsprüfungsprogramme.....	58
11.4	Leistungsbewertung und Bewertung durch Scores bei qualitativen Eignungsprüfungsprogrammen.....	59
Anhang A (normativ) Symbole und Abkürzungen.....		62
Anhang B (informativ) Homogenität und Stabilität von Eignungsprüfungsgegenständen.....		64
B.1	Allgemeines Verfahren zur Homogenitätsprüfung.....	64
B.2	Beurteilungskriterien für eine Homogenitätsprüfung.....	64
B.3	Formeln für die Homogenitätsprüfung.....	67
B.4	Verfahren zur Stabilitätsprüfung.....	69
B.4.1	Allgemeine Betrachtungen zur Stabilitätsprüfung.....	69
B.4.2	Verfahren zur Stabilitätsprüfung im Verlauf einer Eignungsprüfungsrunde.....	70
B.5	Beurteilungskriterium für eine Stabilitätsprüfung.....	71
B.6	Stabilität unter Transportbedingungen.....	72
Anhang C (informativ) Robuste Analyse.....		73
C.1	Allgemeines.....	73

C.2	Einfache ausreißerresistente Schätzfunktionen für den Mittelwert und die Standardabweichung der Grundgesamtheit.....	74
C.2.1	Median .....	74
C.2.2	Skalierter Median der absoluten Abweichungen MAD <sub>e</sub> .....	74
C.2.3	Normierter Interquartilsabstand nIQR.....	74
C.3	Robuste Analyse: Algorithmus A .....	75
C.3.1	Algorithmus A mit iterierter Skala .....	75
C.3.2	Varianten von Algorithmus A .....	76
C.4	Robuste Analyse: Algorithmus S.....	77
C.5	Rechenintensive robuste Schätzfunktionen: Q-Verfahren und Hampel-Schätzer .....	78
C.5.1	Begründung für rechenintensive Schätzfunktionen.....	78
C.5.2	Bestimmung einer robusten Standardabweichung unter Anwendung von Q- und Q <sub>n</sub> -Verfahren .....	79
C.5.3	Bestimmung eines robusten Mittelwerts mit dem Hampel-Schätzer.....	82
C.5.4	Q/Hampel-Verfahren .....	84
C.6	Andere robuste Verfahren.....	85
<b>Anhang D (informativ) Zusätzlicher Leitfaden zu statistischen Verfahren .....</b>		<b>86</b>
D.1	Verfahren für kleine Teilnehmeranzahlen .....	86
D.1.1	Allgemeine Überlegungen .....	86
D.1.2	Verfahren zur Ausreißerererkennung.....	86
D.1.3	Verfahren für lagebezogene Schätzwerte.....	87
D.1.4	Verfahren für Schätzwerte der Streuung.....	87
D.2	Effizienz und Bruchpunkte für robuste Verfahren.....	88
D.2.1	Verschiedene statistische Schätzer (z. B. robuste Verfahren) können anhand von drei Schlüsselmerkmalen verglichen werden: .....	88
D.2.2	Bruchpunkt .....	88
D.2.3	Relative Effizienz .....	89
D.3	Verwendung von Eignungsprüfdaten zur Bewertung der Vergleich- und Wiederholpräzision eines Messverfahrens .....	90
<b>Anhang E (informativ) Erläuternde Beispiele .....</b>		<b>92</b>
E.1	Wirkung von zensierten Werten (siehe 5.5.3.3) .....	92
E.2	Homogenitäts- und Stabilitätsprüfung — Arsen (As) in Schokolade (siehe 6.1).....	93
E.3	Umfassendes Beispiel für Atrazin in Trinkwasser .....	95
E.4	Umfassendes Beispiel für Quecksilber in Tierfutter.....	99
E.5	Referenzwert von einem einzelnen Laboratorium: Los-Angeles-Wert von Zuschlagstoffen (siehe 7.5) .....	103
E.6	Beispiel für ein Bootstrap-Verfahren für coliforme Bakterien in Lebensmittelproben (siehe 7.7.6).....	104
E.7	Vergleich von Referenzwert und Konsens-Mittelwert (siehe 7.8) .....	105
E.8	Bestimmung von Bewertungskriterien durch Erfahrung mit früheren Runden: Toxaphen in Trinkwasser (siehe 8.3) .....	106
E.9	Anhand eines allgemeinen Modells: Horwitz-Gleichung (siehe 8.4) .....	108
E.10	Bestimmung der Leistung aus einem Präzisionsversuch: Bestimmung des Zementgehalts von Festbeton (siehe 8.5).....	108
E.11	Säulendiagramme von standardisierten systematischen Abweichungskomponenten: Antikörperkonzentrationen (siehe 10.4).....	109
E.12	Youden-Diagramm — Antikörper-Konzentrationen (siehe 10.5).....	109
E.13	Diagramm der Wiederholstandardabweichungen: Antikörper-Konzentrationen (siehe 10.6).....	112
E.14	Grafische Verfahren zur Rückverfolgung der Leistung über die Zeit (siehe 10.8) .....	114
E.15	Qualitative Datenanalyse; Beispiel für eine Ordinalgröße: Hautreaktion auf ein Kosmetikum (siehe Abschnitt 11) .....	116
<b>Anhang F (informativ) Beispiel für einen Computercode zur Darstellung und Resampling-Analyse („Bootstrapping“) von Ergebnissen aus Eignungsprüfungen .....</b>		<b>118</b>
<b>Literaturhinweise .....</b>		<b>119</b>

## Bilder

Bild E.1 — Nach Rang geordnete Teilnehmerergebnisse für Atrazin (Daten nach Tabelle E.4) .....	97
Bild E.2 — Histogramm der Teilnehmerergebnisse .....	98
Bild E.3 — Kerndichtediagramm für Teilnehmerergebnisse .....	98
Bild E.4 — Zusammenfassung von robusten Kenngrößen nach Tabelle E.5 .....	99
Bild E.5 — Teilnehmerergebnisse und Unsicherheiten für Ergebnisse in IMEP 111 (Daten nach Tabelle E.6) .....	101
Bild E.6 — Kerndichtediagramm für Teilnehmerergebnisse .....	102
Bild E.7 — Kerndichtediagramm für Teilnehmerergebnisse .....	105
Bild E.8 — Relative Standardabweichung von Teilnehmerergebnissen (%) verglichen mit dem zugewiesenen Wert ( $\mu\text{g/l}$ ) .....	107
Bild E.9 — Teilnehmer-Standardabweichung ( $\mu\text{g/l}$ ) verglichen mit dem zugewiesenen Wert ( $\mu\text{g/l}$ ) .....	108
Bild E.10 — Säulendiagramm der z-Scores (4,0 bis -4,0) für eine Runde eines Eignungsprüfungsprogramms, während der die Teilnehmer Konzentrationen dreier allergenspezifischer IgE-Antikörper bestimmten .....	109
Bild E.11 — Youden-Diagramm der z-Scores nach Tabelle E.10 .....	110
Bild E.12 — Diagramm von Standardabweichungen gegen die Mittelwerte für 25 Teilnehmer (Daten nach Tabelle E.10) .....	113
Bild E.13 — Leistungs-Scores für jede Runde (Daten nach Tabelle E.12) .....	115
Bild E.14 — Leistungs-Scores für verschiedene Niveaus der Messgröße .....	115
Bild E.15 — Säulendiagramm von prozentualen Reaktionen auf zwei Eignungsprüfungsgegenstände für die Hautreizung — # Modus, @ Median .....	117

## Tabellen

Tabelle B.1 — Faktoren $F_1$ und $F_2$ zur Verwendung in der Prüfung auf ausreichende Homogenität .....	66
Tabelle C.1 — Für die robuste Auswertung notwendige Faktoren: Algorithmus S .....	78
Tabelle C.2 — Korrekturfaktor $b_p$ für $2 \leq p \leq 12$ .....	80
Tabelle D.1 — Bruchpunkte für Schätzwerte des Mittelwerts und der Standardabweichung (Anteil von Ausreißern, der zum Versagen des Schätzers führen kann) .....	88
Tabelle D.2 — Relative Effizienz von robusten Schätzern für Mittelwert und Standardabweichung der Grundgesamtheit für normalverteilte Datensätze mit $n = 50$ oder 500 Teilnehmer: .....	89

<b>Tabelle E.1</b>	<b>— Stichprobendatensatz mit gestutzten (&lt;)-Ergebnissen und drei Wahlmöglichkeiten für entsprechende Ergebnisse .....</b>	<b>92</b>
<b>Tabelle E.2</b>	<b>— Homogenitätsdaten für Eignungsprüfungsgegenstände für Arsen in Schokolade .....</b>	<b>93</b>
<b>Tabelle E.3</b>	<b>— Stabilitätsdaten für Eignungsprüfungsgegenstände für Arsen in Schokolade .....</b>	<b>94</b>
<b>Tabelle E.4</b>	<b>— Berechnung des robusten Mittelwerts und der robusten Standardabweichung für Atrazin in Trinkwasser .....</b>	<b>95</b>
<b>Tabelle E.5</b>	<b>— Zusammenfassende Kenngrößen für das Atrazin-Beispiel .....</b>	<b>98</b>
<b>Tabelle E.6</b>	<b>— Eignungsprüfungsergebnisse von 24 Teilnehmern an einer IMEP 111-Untersuchung.....</b>	<b>100</b>
<b>Tabelle E.7</b>	<b>— Leistungskenngrößen durch verschiedene Verfahren .....</b>	<b>102</b>
<b>Tabelle E.8</b>	<b>— Berechnung der mittleren Differenz zwischen einem ZRM und einem Eignungsprüfungsgegenstand und der Standardunsicherheit dieser Differenz.....</b>	<b>103</b>
<b>Tabelle E.9</b>	<b>— Eignungsprüfungsrunden für Toxaphen in Trinkwasser und <math>p \geq 20</math> Ergebnisse.....</b>	<b>106</b>
<b>Tabelle E.10</b>	<b>— Daten und Berechnungen zu Antikörperkonzentrationen für zwei ähnliche Allergene .....</b>	<b>111</b>
<b>Tabelle E.11</b>	<b>— Konzentrationen bestimmter Antikörper in Serum-Eignungsprüfungsgegenständen (vier Wiederholbestimmungen an einem Eignungsprüfungsgegenstand je Teilnehmer).....</b>	<b>112</b>
<b>Tabelle E.12</b>	<b>— <math>P_A</math>-Scores für 5 Runden eines Eignungsprüfungsprogramms, jeweils mit 3 Eignungsprüfungsgegenständen für Kalium in Serum .....</b>	<b>114</b>
<b>Tabelle E.13</b>	<b>— Ergebnisse für zwei Eignungsprüfungsgegenstände, Hautreizung .....</b>	<b>116</b>