

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen

Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision
eines vereinheitlichten Messverfahrens
(ISO 5725-2:1994 einschließlich Technisches Korrigendum 1:2002)

DIN**ISO 5725-2**

ICS 17.020

Mit
DIN ISO 5725-1:1997-11
Ersatz für
die 1997-11 zurückgezogene
Norm DIN ISO 5725:1988-04

Accuracy (trueness and precision) of measurement methods
and results – Part 2: Basic method for the determination of
repeatability and reproducibility of a standard measurement method
(ISO 5725-2:1994 including Technical Corrigendum 1:2002)

Exactitude (justesse et fidélité) des résultats et méthodes
de mesure – Partie 2: Méthode de base pour la détermination
de la répétabilité et de la reproductibilité d'une méthode
de mesure normalisée
(ISO 5725-2:1994, Rectificatif Technique 1:2002 inclus)

Die Internationale Norm ISO 5725-2:1994-12-15 „Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method“ einschließlich Technisches Korrigendum 1:2002 ist unverändert in diese Deutsche Norm übernommen worden.

Nationales Vorwort

Die Internationale Norm ISO 5725 wurde vom Technischen Komitee 69 „Anwendung statistischer Methoden“ der Internationalen Organisation für Normung (ISO) erarbeitet. Die 2. Ausgabe 1986 wurde in einer nachfolgenden 3. Ausgabe 1994 in sechs Teile unterteilt. In diesen sind die Teilaspekte zum Gesamtkomplex „Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen“ einzeln dargestellt. Die deutschen Übersetzungen sind vom Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN erstellt worden. Das Vorwort in jedem der sechs Teile gibt Auskunft über die jeweils behandelten Teilkomplexe.

Die Normen der Reihe DIN ISO 5725 befassen sich mit der Ermittlung und Anwendung von Präzisions- und Richtigkeitsmaßen eines vereinheitlichten Messverfahrens mittels Ringversuchen. Die Genauigkeit eines Messverfahrens wird dadurch beschrieben, dass sowohl ein Richtigkeits- als auch ein Präzisionsmaß gemeinsam angegeben werden. Nach DIN 55350-13 ist Präzision die qualitative Bezeichnung für das Ausmaß der gegenseitigen Annäherung voneinander unabhängiger Ermittlungsergebnisse bei mehrfacher Anwendung eines vereinheitlichten Messverfahrens unter vorgegebenen Bedingungen. Nach derselben Norm ist Richtigkeit die qualitative Bezeichnung für das Ausmaß der Annäherung des Erwartungswertes des Ermittlungsergebnisses an den Bezugswert, wobei dieser je nach Festlegung oder Vereinbarung der wahre oder der richtige Wert sein kann. Der Erwartungswert ist dabei als gedanklich feststellbarer Wert das mittlere Ergebnis bei unablässig wiederholter Anwendung des vereinheitlichten Messverfahrens.

Die Normen der Reihe DIN ISO 5725 sind nur auf kontinuierliche Merkmale anwendbar (siehe 1.2).

Fortsetzung Seite 2 bis 97

Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ)
im DIN Deutsches Institut für Normung e.V.

Als Maß für die Richtigkeit wird die Differenz zwischen einem Schätzwert für den Erwartungswert und dem gewählten Bezugswert benutzt. Die Differenz zwischen Erwartungswert und Bezugswert nennt man auch „bias“. Im Deutschen wird dieses Maß für die Richtigkeit als „systematische Abweichungskomponente“ bezeichnet, um sie von einzelnen systematischen Abweichungen zu unterscheiden.

Maße für die Präzision eines vereinheitlichten Messverfahrens sind im Allgemeinen die Standardabweichungen der unter Wiederholbedingungen, Zwischenbedingungen oder Vergleichbedingungen gewonnenen Messwerte. Aus dem Präzisionsmaß Standardabweichung kann man auch die früher häufig benutzten speziellen Präzisionsmaße Wiederholgrenze r und Vergleichgrenze R als „kritische Differenzbeträge“ berechnen. Diese sind im Vergleich mit den Standardabweichungen um einen angebbaren Faktor größere Beträge, unter denen oder höchstens gleich denen der Betrag der Differenz zweier Einzelergebnisse (gemessen unter den betreffenden Bedingungen) mit einer vorgegebenen Wahrscheinlichkeit P erwartet werden kann. Für diese Wahrscheinlichkeit wird bei Abwesenheit anderer Festlegungen oder Vereinbarungen $P = 95\%$ eingesetzt. Der genannte Faktor hängt von der gewählten oder vorgegebenen Wahrscheinlichkeit ab und hat bei $P = 95\%$ den Zahlenwert 2,8¹⁾. Dem Anwender der Norm (beispielsweise einem Normenausschuss) ist es nach wie vor prinzipiell freigestellt, ob er als Präzisionsmaße eines vereinheitlichten Messverfahrens

- Wiederholstandardabweichung, Zwischenstandardabweichung und Vergleichstandardabweichung oder
- Wiederholgrenze, Zwischengrenze und Vergleichgrenze

angibt. Empfohlen wird, nur die im vorausgehenden ersten Anstrich genannten Größen zu verwenden.

Nicht jeder für ein betrachtetes vereinheitlichtes Messverfahren durchgeführte Ringversuch führt zu den gleichen Werten der Richtigkeitsmaße bzw. der Präzisionsmaße. Diese Ergebnisse sind nämlich Zufallsergebnisse. Überdies können die Anzahl der beteiligten Labors und die Anzahl der Ermittlungen im Labor unterschiedlich sein. Die ermittelten Werte der Richtigkeitsmaße bzw. der Präzisionsmaße sind also Schätzwerte. Für einen Industrie- oder Wirtschaftssektor oder für ein Fachgebiet können für sie jedoch verbindliche Grenzwerte eingeführt werden, falls es erforderlich ist. Das kann beispielsweise aus juristischen Gründen nötig sein.

In den Normen der Reihe DIN ISO 5725 werden außer den eingangs für Präzision und Richtigkeit genannten die folgenden Begriffe benutzt:

- Ringversuch = Versuch unter Beteiligung mehrerer Labors, z. B. zur Ermittlung der Präzision oder der Richtigkeit eines Messverfahrens nach dieser Normenreihe (siehe auch DIN ISO 5725-1:1997-11, 3.22).

ANMERKUNG Ein Ringversuch kann unterschiedliche Zwecke haben. Beispielsweise kann er nur auf Präzisionsmaße oder nur auf Richtigkeitsmaße oder auf beide ausgerichtet sein. Unterbegriffe zum Ringversuch für dessen unterschiedliche Ziele sind nicht erforderlich, werden aber benutzt.

- Labor = Stelle, die Prüfungen oder nur Ermittlungen ausführt.

ANMERKUNG Gelegentlich wird dieser Begriff auch im Sinn des Untersuchungsorts verwendet.

- Vereinheitlichtes Messverfahren: Siehe DIN ISO 5725-1:1997-11, 4.1.

ANMERKUNG Bisher wurde es „festgelegtes Messverfahren“ genannt.

Im Übrigen wird auf die Normen über Begriffe zu Qualitätsmanagement und Statistik der Normenreihen DIN 55350 und DIN 13303 hingewiesen.

Die vorliegende Normenreihe DIN ISO 5725 ist eine Rahmenregelung für zahlreiche Fachgebiete. Sie bedarf bei der Anwendung gegebenenfalls ergänzender Festlegungen, insbesondere zur Anzahl der am Ringversuch beteiligten Labors und zur Anzahl der Merkmalsniveaus und der Wiederholmessungen. Für beides enthält die Normenreihe Empfehlungen. Branchenspezifisch kann eine feste Anzahl oder Mindestanzahl von Labors, Merkmalsniveaus und von Wiederholmessungen pro Merkmalsniveau vorgegeben werden.

Die vorliegende Norm behandelt unter Anwendung der in DIN ISO 5725-1 dargelegten, für alle Normen der Reihe DIN ISO 5725 geltenden allgemeinen Grundlagen und Begriffe, die grundlegenden Methoden für die

¹⁾ Dieser Faktor ergibt sich aus dem auf 2,0 gerundeten Quantil $u_{97,5\%} = 1,96$ sowie aus $\sqrt{n}$ mit $n = 2$. Aus deren Multiplikation folgt, für Zwecke der Schätzung hinreichend genau, der Faktor 2,8.

Ermittlung der beiden Präzisionsmaße eines vereinheitlichten Messverfahrens, bei denen dasselbe vereinheitlichte Messverfahren an identischen Versuchsobjekten benutzt wird, die Untersuchungen aber einmal im selben Labor durch denselben Bearbeiter mit derselben Geräteausrüstung, zum anderen jedoch in verschiedenen Labors durch verschiedene Bearbeiter mit verschiedenen Geräteausrüstungen durchgeführt werden. Dabei ergibt sich im ersten Fall ein Maß für die Wiederholpräzision, im zweiten Fall ein Maß für die Vergleichpräzision.

Maße für die Präzision unter unterschiedlich kombinierten Zwischenbedingungen bezüglich der mitwirkenden Labors, Bearbeiter und Geräteausrüstungen werden in DIN ISO 5725-3 behandelt.

Änderungen

Gegenüber DIN ISO 5725-1:1997-11 und der 1997-11 zurückgezogenen DIN ISO 5725:1988-04 wurden folgende Änderungen vorgenommen:

- Die Norm DIN ISO 5725 wurde in DIN ISO 5725-1 und DIN ISO 5725-2 aufgeteilt. Diese sind terminologisch an die Entwicklung von ISO 3534-1 und ISO 3534-2 angepasst worden.

Frühere Ausgaben

DIN ISO 5725: 1981-11, 1988-04

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN 13303-1:1982-05, *Stochastik — Wahrscheinlichkeitstheorie, Gemeinsame Grundbegriffe der mathematischen und der beschreibenden Statistik — Begriffe und Zeichen.*

DIN 13303-2:1982-11, *Stochastik — Mathematische Statistik — Begriffe und Zeichen.*

DIN 55350-12:1989-03, *Begriffe der Qualitätssicherung und Statistik — Merkmalsbezogene Begriffe.*

DIN 55350-13:1987-07, *Begriffe der Qualitätssicherung und Statistik — Begriffe zur Genauigkeit von Ermittlungsverfahren und Ermittlungsergebnissen.*

DIN 55350-14:1985-12, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Probenahme.*

DIN 55350-15:1986-02, *Begriffe der Qualitätssicherung und Statistik — Begriffe zu Mustern.*

DIN 55350-17:1988-08, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Qualitätsprüfungsarten.*

DIN 55350-21:1982-05, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Zufallsgrößen und Wahrscheinlichkeitsverteilungen.*

DIN 55350-22:1987-02, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Spezielle Wahrscheinlichkeitsverteilungen.*

DIN 55350-23:1983-04, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Beschreibende Statistik.*

DIN 55350-24:1982-11, *Begriffe der Qualitätssicherung und Statistik — Begriffe der Statistik — Schließende Statistik.*

DIN EN ISO 9000:2000-12, *Qualitätsmanagementsysteme — Grundlagen und Begriffe.*

Deutsche Übersetzung

Genauigkeit (Richtigkeit und Präzision) von Messverfahren und
Messergebnissen –

Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und
Vergleichpräzision eines vereinheitlichten Messverfahrens

Inhalt

	Seite
1 Anwendungsbereich	12
2 Normative Verweisungen	14
3 Begriffe	14
4 Schätzwerte für die Parameter des Grundmodells	14
5 Forderungen an einen Präzisionsversuch	14
5.1 Gestaltung des Präzisionsversuchs	14
5.2 Auswahl der Labors	16
5.3 Vorbereitung des Materials	16
6 Am Präzisionsversuch beteiligtes Personal	20
6.1 Versuchsgremium	20
6.2 Statistische Aufgaben	20
6.3 Führungsaufgaben	20
6.4 Laborverantwortliche	20
6.5 Bearbeiter	22
7 Statistische Auswertung des Präzisionsversuchs	22
7.1 Einleitende Überlegungen	22
7.2 Tabellarische Zusammenstellung der Ergebnisse sowie angewendete Darstellungsweise	24
7.3 Ergebnisprüfung bezüglich Vereinbarkeit und Ausreißer	28
7.4 Errechnung des Gesamtmittelwerts und der Varianzen	36
7.5 Feststellen einer funktionalen Beziehung zwischen Präzisionswerten und dem arithmetischen Gesamtmittelwert m	40
7.6 Statistische Auswertung als schrittweises Verfahren	46
7.7 Der Bericht an das Versuchsgremium sowie dessen Entscheidungen	52
8 Statistische Tabellen	54

Accuracy (trueness and precision) of measurement methods and results —

Part 2:

Basic method for the determination of repeatability and reproducibility of a standard measurement method

Contents

	Page
1 Scope	13
2 Normative references	15
3 Definitions	15
4 Estimates of the parameters in the basic model	15
5 Requirements for a precision experiment	15
5.1 Layout of the experiment	15
5.2 Recruitment of the laboratories	17
5.3 Preparation of the materials	17
6 Personnel involved in a precision experiment	21
6.1 Panel	21
6.2 Statistical functions	21
6.3 Executive functions	21
6.4 Supervisors	21
6.5 Operators	23
7 Statistical analysis of a precision experiment	23
7.1 Preliminary considerations	23
7.2 Tabulation of the results and notation used	25
7.3 Scrutiny of results for consistency and outliers	29
7.4 Calculation of the general mean and variances	37
7.5 Establishing a functional relationship between precision values and the mean level m	41
7.6 Statistical analysis as a step-by-step procedure	47
7.7 The report to, and the decisions to be taken by, the panel	53
8 Statistical tables	55

	Seite
Anhang A	Formelzeichen und Abkürzungen in ISO 5725 62
Anhang B	Beispiele für die statistische Auswertung von Präzisionsversuchen 66
B.1	Beispiel 1: Ermittlung des Schwefelgehalts von Kohle (Mehrere Merkmalsniveaus ohne fehlende Daten oder Ausreißer).....66
B.2	Beispiel 2: Erweichungspunkt von Pech (Mehrere Merkmalsniveaus mit fehlenden Daten).....76
B.3	Beispiel 3: Thermometrische Titration von Kreosotöl (Mehrere Merkmalsniveaus mit Ausreißern).....84
Anhang C	Literaturhinweise.....96

Annexes

A	Symbols and abbreviations used in ISO 5725	63
B	Examples of the statistical analysis of precision experiments	67
B.1	Example 1: Determination of the sulfur content of coal (Several levels with no missing or outlying data)	67
B.2	Example 2: Softening point of pitch (Several levels with missing data)	77
B.3	Example 3: Thermometric titration of creosote oil (Several levels with outlying data)	85
C	Bibliography	97

Vorwort

Die ISO (Internationale Organisation für Normung) ist die weltweite Vereinigung nationaler Normungsinstitute (ISO-Mitgliedskörperschaften). Die Erarbeitung Internationaler Normen obliegt den Technischen Komitees der ISO. Jede Mitgliedskörperschaft, die sich für ein Thema interessiert, für das ein Technisches Komitee eingesetzt wurde, ist berechtigt, in diesem Komitee mitzuarbeiten. Internationale (staatliche und nichtstaatliche) Organisationen, die mit der ISO in Verbindung stehen, sind an den Arbeiten ebenfalls beteiligt. Die ISO arbeitet in allen Angelegenheiten der elektrotechnischen Normung eng mit der Internationalen Elektrotechnischen Kommission (IEC) zusammen.

Die von den Technischen Komitees verabschiedeten internationalen Norm-Entwürfe werden den Mitgliedskörperschaften zur Abstimmung vorgelegt. Die Veröffentlichung als Internationale Norm erfordert Zustimmung von mindestens 75 % der abstimmenden Mitgliedskörperschaften.

Die Internationale Norm ISO 5725-2 wurde durch das Technische Komitee ISO/TC 69, *Anwendung statistischer Methoden*, Unterkomitee SC 6, *Messmethoden und Messergebnisse*, erarbeitet.

ISO 5725 besteht unter dem Haupttitel „*Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen*“ aus den folgenden Teilen:

- Teil 1: Allgemeine Grundlagen und Begriffe
- Teil 2: Grundlegende Methode für die Ermittlung der Wiederhol- und Vergleichpräzision eines vereinheitlichten Messverfahrens
- Teil 3: Präzisionsmaße eines vereinheitlichten Messverfahrens unter Zwischenbedingungen
- Teil 4: Grundlegende Methoden für die Ermittlung der Richtigkeit eines vereinheitlichten Messverfahrens
- Teil 5: Alternative Methoden für die Ermittlung der Präzision eines vereinheitlichten Messverfahrens
- Teil 6: Anwendung von Genauigkeitswerten in der Praxis

Die Teile 1 bis 6 von ISO 5725 ersetzen gemeinsam die Norm ISO 5725:1986 und machen diese ungültig, welche erweitert wurde, um Richtigkeit (ergänzend zu Präzision) sowie Zwischenbedingungen (ergänzend zu Wiederhol- und Vergleichbedingungen) abzudecken.

Der Anhang A ist normativer Bestandteil dieses Teils von ISO 5725, die Anhänge B und C sind informativ.

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.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 5725-2 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*, Subcommittee SC 6, *Measurement methods and results*.

ISO 5725 consists of the following parts, under the general title *Accuracy (trueness and precision) of measurement methods and results*:

- *Part 1: General principles and definitions*
- *Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*
- *Part 3: Intermediate measures of the precision of a standard measurement method*
- *Part 4: Basic methods for the determination of the trueness of a standard measurement method*
- *Part 5: Alternative methods for the determination of the precision of a standard measurement method*
- *Part 6: Use in practice of accuracy values*

Parts 1 to 6 of ISO 5725 together cancel and replace ISO 5725:1986, which has been extended to cover trueness (in addition to precision) and intermediate precision conditions (in addition to repeatability and reproducibility conditions).

Annex A forms an integral part of this part of ISO 5725. Annexes B and C are for information only.

Einleitung

0.1 ISO 5725 benutzt die beiden Ausdrücke „Richtigkeit“ und „Präzision“, um die Genauigkeit eines Messverfahrens zu beschreiben. „Richtigkeit“ bezieht sich auf das Ausmaß der Annäherung zwischen dem arithmetischen Mittelwert einer großen Anzahl von Ermittlungsergebnissen und dem wahren oder dem richtigen Wert. „Präzision“ bezieht sich auf das Ausmaß der gegenseitigen Annäherung von Ermittlungsergebnissen.

0.2 Allgemeine Überlegungen zu diesen Größen sind in ISO 5725-1 angegeben. Sie werden deshalb in diesem Teil von ISO 5725 nicht wiederholt. ISO 5725-1 erläutert die grundlegenden Definitionen und allgemeinen Prinzipien und sollte deshalb zusammen mit allen anderen Teilen von ISO 5725 gelesen werden, also auch zusammen mit diesem Teil 2.

0.3 Dieser Teil von ISO 5725 befasst sich nur mit Schätzungen mittels der Wiederholstandardabweichung und der Vergleichstandardabweichung. Obwohl unter speziellen Umständen andere Versuchstypen (wie etwa der Versuch mit Doppelniveauproben) für die Schätzung der Präzision angewendet werden, sind solche in diesem Teil von ISO 5725 nicht behandelt. Sie sind vielmehr Gegenstand von ISO 5725-5. Dieser Teil von ISO 5725 beschäftigt sich auch nicht mit anderen Präzisionsmaßen, die zwischen den beiden hauptsächlichen Präzisionsmaßen liegen. Präzisionsmaße unter Zwischenbedingungen sind Gegenstand von ISO 5725-3.

0.4 Unter speziellen Umständen können Ergebnisse eines Versuchs zur Schätzung der Präzision auch zur Schätzung der Richtigkeit verwendet werden. Die Schätzung der Richtigkeit ist in diesem Teil von ISO 5725 nicht behandelt. Alle Gesichtspunkte zur Schätzung der Richtigkeit sind Gegenstand von ISO 5725-4.

Introduction

0.1 ISO 5725 uses two terms "trueness" and "precision" to describe the accuracy of a measurement method. "Trueness" refers to the closeness of agreement between the arithmetic mean of a large number of test results and the true or accepted reference value. "Precision" refers to the closeness of agreement between test results.

0.2 General consideration of these quantities is given in ISO 5725-1 and so is not repeated in this part of ISO 5725. ISO 5725-1 should be read in conjunction with all other parts of ISO 5725, including this part, because it gives the underlying definitions and general principles.

0.3 This part of ISO 5725 is concerned solely with estimating by means of the repeatability standard deviation and reproducibility standard deviation. Although other types of experiment (such as the split-level experiment) are used in certain circumstances for the estimation of precision, they are not dealt with in this part of ISO 5725 but rather are the subject of ISO 5725-5. Nor does this part of ISO 5725 consider any other measures of precision intermediate between the two principal measures; those are the subject of ISO 5725-3.

0.4 In certain circumstances, the data obtained from an experiment carried out to estimate precision are used also to estimate trueness. The estimation of trueness is not considered in this part of ISO 5725; all aspects of the estimation of trueness are the subject of ISO 5725-4.

1 Anwendungsbereich

1.1 Dieser Teil von ISO 5725

- erläutert ausführlich die allgemeinen Prinzipien, die zu beachten sind, wenn man Versuche zur zahlenmäßigen Schätzung der Präzision eines (vereinheitlichten) Messverfahrens mittels Ringversuch entwirft;
- vermittelt im Einzelnen eine praktische Beschreibung der Grundmethode zur formellen Anwendung der Schätzung der Präzision von (vereinheitlichten) Messverfahren;
- vermittelt eine Anleitung für alle Personen, die mit dem Planen, dem Ausführen oder dem Auswerten der Ergebnisse der Versuche zur Schätzung der Präzision befasst sind.

ANMERKUNG 1 Abwandlungen zu dieser Grundmethode für besondere Zwecke sind in anderen Teilen von ISO 5725 angegeben.

Anhang B vermittelt praktische Beispiele zur Schätzung der Präzision (vereinheitlichter) Messverfahren mittels Ringversuchen.

1.2 Dieser Teil von ISO 5725 befasst sich ausschließlich mit Messverfahren für kontinuierliche Merkmale, die als Ermittlungsergebnis einen einzigen Wert liefern, auch wenn dieser einzige Wert ein Rechenergebnis aus einer Reihe von Beobachtungen sein kann.

1.3 Dieser Teil von ISO 5725 geht davon aus, dass beim Entwurf und bei der Durchführung des Präzisionsversuchs alle Prinzipien beachtet werden, die in ISO 5725-1 niedergelegt sind. Die Grundmethode wendet die gleiche Anzahl von Ermittlungsergebnissen in jedem Labor an, wobei jedes Labor die gleichen Merkmalsniveaus der Proben untersucht. Man nennt das einen „ausgewogenen Versuch auf einheitlichem Merkmalsniveau“. Die Grundmethode ist anwendbar auf Verfahren, die ver-

einheitlicht worden sind und in einer Anzahl von Labors regelmäßig angewendet werden.

ANMERKUNG 2 Ausgearbeitete Beispiele sind angegeben, um ausgewogene Ermittlungsergebnisse mit jeweils gleicher Anzahl (je Zelle) darzulegen. Allerdings wird in einem der Beispiele von einer schwankenden Anzahl von Ermittlungsergebnissen je Zelle berichtet (unausgewogener Versuchsplan). In einem anderen Beispiel fehlen einige Daten. Das liegt daran, dass sich ein ausgewogen geplanter Versuch in einen nicht ausgewogenen umwandeln kann. Fastausreißer und Ausreißer sind ebenfalls betrachtet.

1.4 Das statistische Modell nach ISO 5725-1:1994, Abschnitt 5, ist als eine geeignete Grundlage für die Deutung und Auswertung solcher Ermittlungsergebnisse übernommen, deren Verteilung näherungsweise eine Normalverteilung ist.

1.5 Die Grundmethode nach Beschreibung in diesem Teil von ISO 5725 wird (üblicherweise) die Präzision eines (vereinheitlichten) Messverfahrens schätzen,

- a) wenn die Ermittlung der Wiederhol- und der Vergleichstandardabweichungen gefordert ist, wie sie in ISO 5725-1 definiert sind;
- b) wenn die anzuwendenden Materialien homogen sind, oder wenn die Auswirkungen einer Inhomogenität in die Präzisionswerte einbezogen werden können; und

1 Scope

1.1 This part of ISO 5725

- amplifies the general principles to be observed in designing experiments for the numerical estimation of the precision of measurement methods by means of a collaborative interlaboratory experiment;
- provides a detailed practical description of the basic method for routine use in estimating the precision of measurement methods;
- provides guidance to all personnel concerned with designing, performing or analysing the results of the tests for estimating precision.

NOTE 1 Modifications to this basic method for particular purposes are given in other parts of ISO 5725.

Annex B provides practical examples of estimating the precision of measurement methods by experiment.

1.2 This part of ISO 5725 is concerned exclusively with measurement methods which yield measurements on a continuous scale and give a single value as the test result, although this single value may be the outcome of a calculation from a set of observations.

1.3 It assumes that in the design and performance of the precision experiment, all the principles as laid down in ISO 5725-1 have been observed. The basic method uses the same number of test results in each laboratory, with each laboratory analysing the same levels of test sample; i.e. a balanced uniform-level experiment. The basic method applies to procedures that have been standardized and are in regular use in a number of laboratories.

NOTE 2 Worked examples are given to demonstrate balanced uniform sets of test results, although in one example a variable number of replicates per cell were reported (unbalanced design) and in another some data were missing. This is because an experiment designed to be balanced can turn out to be unbalanced. Stragglers and outliers are also considered.

1.4 The statistical model of clause 5 of ISO 5725-1:1994 is accepted as a suitable basis for the interpretation and analysis of the test results, the distribution of which is approximately normal.

1.5 The basic method, as described in this part of ISO 5725, will (usually) estimate the precision of a measurement method:

- a) when it is required to determine the repeatability and reproducibility standard deviations as defined in ISO 5725-1;
- b) when the materials to be used are homogeneous, or when the effects of heterogeneity can be included in the precision values; and

- c) wenn die Anwendung eines ausgewogenen Versuchsplans auf einheitlichem Merkmalsniveau annehmbar ist.

1.6 Derselbe Ansatz kann angewendet werden für eine vorläufige Schätzung der Präzision für Messverfahren, die noch keine Vereinheitlichung erreicht haben oder noch nicht ständig angewendet werden.

2 Normative Verweisungen

Die folgenden normativen Dokumente enthalten Festlegungen, die durch Verweisung in diesem Text Bestandteil dieses Teils von ISO 5725 sind. Zum Zeitpunkt der Veröffentlichung dieser Internationalen Norm waren die angegebenen Ausgaben gültig. Alle Normen unterliegen der Überarbeitung. Vertragspartner, deren Vereinbarungen auf diesem Teil von ISO 5725 basieren, werden gebeten, die Möglichkeit zu prüfen, ob die jeweils neuesten Ausgaben der im Folgenden genannten Normen angewendet werden können. Die Mitglieder von IEC und ISO führen Verzeichnisse der gegenwärtig gültigen Internationalen Normen.

ISO 3534-1:1993, *Statistik — Begriffe und Formelzeichen — Teil 1: Wahrscheinlichkeit und allgemeine statistische Begriffe*.

ISO 5725-1:1994, *Genauigkeit (Richtigkeit und Präzision) von Messverfahren und Messergebnissen — Teil 1: Allgemeine Grundlagen und Begriffe*.

3 Begriffe

Für diesen Teil von ISO 5725 gelten die in ISO 3534-1 und ISO 5725-1 angegebenen Begriffe.

Die in (der Normenreihe) ISO 5725 angewendeten Formelzeichen sind im Anhang A angegeben.

4 Schätzwerte für die Parameter des Grundmodells

4.1 Die in diesem Teil von ISO 5725 angegebenen Verfahren gründen sich auf das in ISO 5725-1:1994, Abschnitt 5, angegebene und in ISO 5725-1:1994, 1.2, weiter ausgearbeitete statistische Modell. Vor allem gründen sich diese Verfahren auf die Gleichungen (2) bis (6) in ISO 5725-1:1994, Abschnitt 5.

Das Modell lautet

$$y = m + B + e.$$

Dabei ist für das jeweils untersuchte Material

- m der Gesamtmittelwert (als Schätzwert für den Erwartungswert);
- B die systematische Abweichungskomponente des Labors unter Wiederholbedingungen;
- e die zufällige Abweichung, die bei jeder Messung unter Wiederholbedingungen vorkommt.

4.2 Die Gleichungen (2) bis (6) in ISO 5725-1:1994, Abschnitt 5, sind mittels der wahren Standardabweichungen der betrachteten Grundgesamtheiten ausgedrückt. In der Praxis sind die wirklichen Werte dieser (wahren) Standardabweichungen unbekannt. Deshalb müssen Schätzwerte für die Präzisionswerte aufgrund einer vergleichsweise kleinen Stichprobe aller in Frage kommenden Labors ermittelt werden, und innerhalb dieser Labors aufgrund einer kleinen Stichprobe aller in Frage kommenden Ermittlungsergebnisse.

4.3 In der statistischen Praxis, wo der wahre Wert einer Standardabweichung, σ , unbekannt ist und durch einen Stichprobenschätzwert ersetzt wird, wird demzufolge das Formelzeichen σ durch das Formelzeichen s ersetzt, um deutlich zu machen, dass dies ein Schätzwert ist. Das muss in jeder der Gleichungen (2) bis (6) in ISO 5725-1:1994 geschehen, woraus folgt:

s_L^2 ist der Schätzwert für die „Varianz zwischen den Labors“;

s_W^2 ist der Schätzwert für die „Varianz innerhalb eines Labors“;

s_r^2 ist der arithmetische Mittelwert von s_W^2 und zugleich Schätzwert für die Wiederholvarianz; dieser arithmetische Mittelwert ist über alle Labors zu erstrecken, die sich am Ringversuch beteiligen und nach dem Ausscheiden von Ausreißern verbleiben;

s_R^2 ist der Schätzwert für die Vergleichvarianz gemäß der Gleichung

$$s_R^2 = s_L^2 + s_r^2 \quad \dots (1)$$

5 Forderungen an einen Präzisionsversuch

5.1 Gestaltung des Präzisionsversuchs

5.1.1 Bei der für die Grundmethode angewendeten Gestaltung werden Stichproben aus q Materiallosen, die q unterschiedliche Merkmalsniveaus für den Versuch repräsentieren, an p Labors gesandt, deren jedes sich für jedes der q Merkmalsniveaus genau n wiederholte Ermittlungsergebnisse unter Wiederholbedingungen verschafft. Diese Art von Versuch wird

„Ausgewogener Versuch auf einheitlichem Merkmalsniveau“

genannt.

- c) when the use of a balanced uniform-level layout is acceptable.

1.6 The same approach can be used to make a preliminary estimate of precision for measurement methods which have not reached standardization or are not in routine use.

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this part of ISO 5725. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this part of ISO 5725 are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3534-1:1993, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*.

ISO 5725-1:1994, *Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions*.

3 Definitions

For the purposes of this part of ISO 5725, the definitions given in ISO 3534-1 and in ISO 5725-1 apply.

The symbols used in ISO 5725 are given in annex A.

4 Estimates of the parameters in the basic model

4.1 The procedures given in this part of ISO 5725 are based on the statistical model given in clause 5 of ISO 5725-1:1994 and elaborated upon in subclause 1.2 of ISO 5725-1:1994. In particular, these procedures are based on equations (2) to (6) of clause 5 of ISO 5725-1:1994.

The model is

$$y = m + B + e$$

where, for the particular material tested,

m is the general mean (expectation);

B is the laboratory component of bias under repeatability conditions;

e is the random error occurring in every measurement under repeatability conditions.

4.2 Equations (2) to (6) of ISO 5725-1:1994, clause 5 are expressed in terms of the true standard deviations of the populations considered. In practice, the exact values of these standard deviations are not known, and estimates of precision values must be made from a relatively small sample of all the possible laboratories, and within those laboratories from a small sample of all the possible test results.

4.3 In statistical practice, where the true value of a standard deviation, σ , is not known and is replaced by an estimate based upon a sample, then the symbol σ is replaced by s to denote that it is an estimate. This has to be done in each of the equations (2) to (6) of ISO 5725-1:1994, giving:

s_L^2 is the estimate of the between-laboratory variance;

s_W^2 is the estimate of the within-laboratory variance;

s_r^2 is the arithmetic mean of s_W^2 and is the estimate of the repeatability variance; this arithmetic mean is taken over all those laboratories taking part in the accuracy experiment which remain after outliers have been excluded;

s_R^2 is the estimate of the reproducibility variance:

$$s_R^2 = s_L^2 + s_r^2 \quad \dots (1)$$

5 Requirements for a precision experiment

5.1 Layout of the experiment

5.1.1 In the layout used in the basic method, samples from q batches of materials, representing q different levels of the test, are sent to p laboratories which each obtain exactly n replicate test results under repeatability conditions at each of the q levels. This type of experiment is called a balanced uniform-level experiment.

5.1.2 Die Ausführung dieser Messungen muss systematisch geplant werden. Dazu müssen folgende Anweisungen gegeben werden:

- a) Jedes vorausgehende Prüfen von Messgeräten muss so durchgeführt werden, wie es in dem vereinheitlichten Messverfahren festgelegt ist.
- b) Jede Gruppe von n Messungen, die ein Merkmalsniveau betrifft, muss unter Wiederholbedingungen ausgeführt werden, das heißt in einer kurzen Zeitspanne, durch denselben Bearbeiter, und ohne irgendwelche zwischengeschaltete erneute Kalibrierung des Messgeräts; es sei denn, dies ist ein Bestandteil der Ausführung einer Messung.
- c) Es ist unbedingt erforderlich, dass jeder der n Versuche unter Wiederholbedingungen unabhängig voneinander ausgeführt wird, so als ob es n Versuche an unterschiedlichen Materialien wären. In der Regel wird allerdings der Bearbeiter oder die Bearbeiterin wissen, dass er (sie) identisches Material prüft. Deshalb sollte in den Anleitungen besonders der Gesichtspunkt hervorgehoben werden, dass es der ganze Zweck des Versuchs ist, zu ermitteln, welche Unterschiede in den Ergebnissen während der Versuchsdurchführung vorkommen können. Wenn trotz dieser Hinweise zu befürchten ist, dass vorausgehende Ermittlungsergebnisse nachfolgende Ermittlungsergebnisse und damit die Wiederholvarianz beeinflussen könnten, sollte überlegt werden, ob auf jedem der q Merkmalsniveaus n verschiedene Proben verwendet werden sollten, die in einer Weise codiert sind, dass der Bearbeiter nicht weiß, was die Proben für ein gegebenes Merkmalsniveau sind. Ein solches Verfahren könnte allerdings dennoch Probleme bezüglich der Sicherstellung von tatsächlich zutreffenden Wiederholbedingungen zwischen den Proben verursachen. Das dürfte nur dann möglich sein, wenn die Messungen in einer Art und Weise ausgeführt werden, dass alle q mal n Messungen innerhalb einer kurzen Zeitspanne ausgeführt werden können.
- d) Es ist nicht unbedingt nötig, dass jede der q Gruppen von je n Messungen alle genau in einer kurzen Zeitspanne ausgeführt werden. Unterschiedliche Gruppen von Messungen können an verschiedenen Tagen ausgeführt werden.
- e) Messungen aller q Merkmalsniveaus müssen durch ein und denselben Bearbeiter ausgeführt werden, und zusätzlich müssen die n Messungen auf einem gegebenen Merkmalsniveau durchgehend mit demselben Messgerät ausgeführt werden.
- f) Sofern im Verlauf der Messungen ein Bearbeiter ausfällt, kann ein anderer Bearbeiter die Messungen zu Ende führen, vorausgesetzt, dass der Wechsel nicht innerhalb ein und derselben Gruppe von n Messungen auf einem Merkmalsniveau vorkommt, sondern nur zwischen zwei der q Gruppen. Über einen solchen Wechsel muss mit den Ergebnissen berichtet werden.
- g) Für die vollständige Durchführung aller Messungen muss eine Höchstzeitspanne vorgegeben werden. Das kann nötig sein, um die Zeitspanne zu begrenzen, die zwischen dem Tag des Empfangs der Pro-

ben und dem Tag des Abschlusses der Messungen verstreichen darf.

- h) Alle Proben müssen mit dem Namen des Versuchs und einer Kennzeichnung der Probe eindeutig etikettiert sein.

5.1.3 In 5.1.2 und an anderer Stelle in diesem Teil von ISO 5725 ist der Bearbeiter erwähnt. Für einige Messungen kann es in Wirklichkeit ein Team von Bearbeitern sein, von denen jeder einen speziellen Teil des Verfahrens ausführt. In einem solchen Fall muss das Team als „der Bearbeiter“ betrachtet werden, und irgendeine Änderung innerhalb des Teams als Bereitstellung eines anderen „Bearbeiters“.

5.1.4 In der Geschäftspraxis können Ergebnisse ziemlich grob gerundet werden. In einem Präzisionsversuch hingegen müssen Ergebnisse mindestens mit einer Stelle mehr als in dem vereinheitlichten Messverfahren verlangt aufgeschrieben werden. Wenn das vereinheitlichte Messverfahren die Anzahl Stellen nicht festlegt, darf die Rundung nicht gröber sein als die Hälfte des Schätzwertes der Wiederholstandardabweichung. Sollte die Präzision vom Merkmalsniveau m abhängen (siehe Gleichung in 4.1), können für unterschiedliche Merkmalsniveaus unterschiedliche Rundungsgrade nötig sein.

5.2 Auswahl der Labors

5.2.1 Die allgemeinen Prinzipien bezüglich der Auswahl der Labors, die an einem Ringversuch teilnehmen sollen, sind in ISO 5725-1:1994, 6.3, angegeben. Bei der Anwerbung der Mitarbeit der erforderlichen Anzahl von Labors müssen deren Verantwortlichkeiten zweifelsfrei festgelegt werden. Ein Beispiel für einen geeigneten Anwerbungs-Fragebogen zeigt Bild 1.

5.2.2 Für den Zweck dieses Teils von ISO 5725 wird ein „Labor“ als eine Kombination aus Bearbeiter, dem Messgerät und der Prüfstelle betrachtet. Eine einzelne Prüfstelle (oder ein Labor im herkömmlichen Sinn) kann deshalb mehrere „Labors“ verkörpern, sofern es mehrere Bearbeiter mit voneinander unabhängigen Messgeräten bereitstellen kann, mit denen die Arbeit unabhängig voneinander durchgeführt wird.

5.3 Vorbereitung des Materials

5.3.1 Eine Besprechung der Gesichtspunkte, die betrachtet werden müssen, wenn Material für die Anwendung in einem Präzisionsversuch ausgewählt wird, ist in ISO 5725-1:1994, 6.4, zu finden.

5.3.2 Bei der Entscheidung über die Mengen des bereitzustellenden Materials müssen versehentliche Mengenverluste oder Irrtümer bei der Beschaffung einiger Prüfergebnisse berücksichtigt werden, was den Einsatz von zusätzlichem Material nötig machen kann. Die Menge des vorbereiteten Materials muss für den Versuch ausreichend sein und einen angemessenen Reserve-Vorrat erlauben.

5.1.2 The performance of these measurements shall be organized and instructions issued as follows.

- a) Any preliminary checking of equipment shall be as specified in the standard method.
- b) Each group of n measurements belonging to one level shall be carried out under repeatability conditions, i.e. within a short interval of time and by the same operator, and without any intermediate recalibration of the apparatus unless this is an integral part of performing a measurement.
- c) It is essential that a group of n tests under repeatability conditions be performed independently as if they were n tests on different materials. As a rule, however, the operator will know that he/she is testing identical material, but the point should be stressed in the instructions that the whole purpose of the experiment is to determine what differences in results can occur in actual testing. If it is feared that, despite this warning, previous results may influence subsequent test results and thus the repeatability variance, it should be considered whether to use n separate samples at each of the q levels, coded in such a way that the operator will not know which are the replicates for a given level. However, such a procedure could cause problems in ensuring that repeatability conditions will apply between replicates. This would only be possible if the measurements were of such a nature that all the qn measurements could be performed within a short interval of time.
- d) It is not essential that all the q groups of n measurements each be performed strictly within a short interval; different groups of measurements may be carried out on different days.
- e) Measurements of all q levels shall be performed by one and the same operator and, in addition, the n measurements at a given level shall be performed using the same equipment throughout.
- f) If in the course of the measurements an operator should become unavailable, another operator may complete the measurements, provided that the change does not occur within a group of n measurements at one level but only occurs between two of the q groups. Any such change shall be reported with the results.
- g) A time limit shall be given within which all measurements shall be completed. This may be necessary to limit the time allowed to elapse be-

tween the day the samples are received and the day the measurements are performed.

- h) All samples shall be clearly labelled with the name of the experiment and a sample identification.

5.1.3 In 5.1.2 and elsewhere in this part of ISO 5725, reference is made to the operator. For some measurements, there may in fact be a team of operators, each of whom performs some specific part of the procedure. In such a case, the team shall be regarded as "the operator" and any change in the team shall be regarded as providing a different "operator".

5.1.4 In commercial practice, the test results may be rounded rather crudely, but in a precision experiment test results shall be reported to at least one more digit than specified in the standard method. If the method does not specify the number of digits, the rounding shall not be coarser than half the repeatability standard deviation estimate. When precision may depend on the level m , different degrees of rounding may be needed for different levels.

5.2 Recruitment of the laboratories

5.2.1 The general principles regarding recruitment of the laboratories to participate in an interlaboratory experiment are given in 6.3 of ISO 5725-1:1994. In enlisting the cooperation of the requisite number of laboratories, their responsibilities shall be clearly stated. An example of a suitable enlistment questionnaire is given in figure 1.

5.2.2 For the purposes of this part of ISO 5725, a "laboratory" is considered to be a combination of the operator, the equipment and the test site. One test site (or laboratory in the conventional sense) may thus produce several "laboratories" if it can provide several operators each with independent sets of equipment and situations in which to perform the work.

5.3 Preparation of the materials

5.3.1 A discussion of the points that need to be considered when selecting materials for use in a precision experiment is given in 6.4 of ISO 5725-1:1994.

5.3.2 When deciding on the quantities of material to be provided, allowance shall be made for accidental spillage or errors in obtaining some test results which may necessitate using extra material. The amount of material prepared shall be sufficient to cover the experiment and allow an adequate stock in reserve.

Fragebogen für einen Ringversuch	
Titel des vereinheitlichten Messverfahrens (siehe beigefügte Kopie):	
1 Unser Labor hat die Absicht, sich an einem Präzisionsversuch für dieses vereinheitlichte Messverfahren zu beteiligen.	
Ja ☐	Nein ☐ (Zutreffendes ankreuzen)
2 Als Teilnehmer wissen wir, dass	
a) alle erforderlichen Geräte, Chemikalien und andere durch das Verfahren festgelegten Forderungen in unserem Labor verfügbar sein müssen, sobald das Programm beginnt;	
b) die festgelegten Forderungen bezüglich „Zeit“, wie das Datum des Beginns, das Anfordern von Prüfproben sowie das Datum für die Beendigung des Programms, strikt erfüllt werden müssen;	
c) wir uns an das vereinheitlichte Messverfahren genau zu halten haben;	
d) Proben genau nach den Anleitungen behandelt werden müssen;	
e) ein qualifizierter Bearbeiter die Messungen ausführen muss.	
Wir haben das vereinheitlichte Messverfahren studiert sowie eine faire Abschätzung unserer Fähigkeiten und Möglichkeiten vorgenommen. Wir glauben, dass wir angemessen für die gemeinsame Prüfung dieses Messverfahrens vorbereitet sein werden.	
3 <u>Kommentare</u>	
	Unterschrift: Unternehmen oder Labor:

Bild 1 — Anwerbungs-Fragebogen für einen Ringversuch

5.3.3 Es sollte überlegt werden, ob es für einige Labors wünschenswert ist, zur Eingewöhnung in das (vereinheitlichte) Messverfahren noch vor Erarbeitung des offiziellen Ermittlungsergebnisses einige vorläufige Ermittlungsergebnisse zu gewinnen, und ob gegebenenfalls für diesen Zweck zusätzliches Material (nicht aber Proben für den Präzisionsversuch) bereitgestellt werden sollte.

5.3.4 Wenn ein Material homogenisiert werden muss, dann muss das in der für dieses Material zweckmäßigsten Art und Weise geschehen. Wenn das zu prüfende Material nicht homogen ist, ist es wichtig, die Proben so vorzubereiten, wie es in dem (vereinheitlichten) Messverfahren festgelegt ist, vorzugsweise beginnend mit einem Los handelsüblichen Materials für jedes Merkmalsniveau. Im Fall von instabilem Material müssen spezielle Anleitungen für Lagerung und Behandlung festgelegt werden.

5.3.5 Für die Proben zu jedem Merkmalsniveau müssen n getrennte Behälter für jedes Labor verwendet werden, sofern irgendeine Gefahr der Beeinträchtigung des Materials besteht, sobald die Behälter geöffnet worden sind (beispielsweise durch Oxydation, durch Verlust flüchtiger Komponenten oder bei hygroskopischem Material).

Im Fall von instabilem Material müssen spezielle Anweisungen für Lagerung und Behandlung festgelegt werden. Vorsichtsmaßnahmen können nötig werden, um sicherzustellen, dass Proben bis zum Zeitpunkt der Durchführung der Messungen unverändert bleiben. Besteht das zu prüfende Material aus einem Gemisch von Pulvern unterschiedlicher relativer Dichte oder unterschiedlicher Korngröße, ist Vorsicht geboten im Hinblick auf Entmischung infolge von Erschütterungen, beispielsweise während eines Transports. Wenn eine Reaktion mit der Atmosphäre erwartet werden kann, können die Proben in Ampullen versiegelt werden, die entweder evakuiert oder mit einem Inertgas gefüllt sind. Für verderbliches Material, wie etwa Lebensmittel- oder Blutproben, kann es notwendig werden, sie in tiefgefrorenem Zustand mit ins Einzelne gehenden Anweisungen für das Verfahren des Auftauens an die beteiligten Labors zu senden.

Questionnaire for interlaboratory study	
Title of measurement method (copy attached)	
1. Our laboratory is willing to participate in the precision experiment for this standard measurement method.	
YES ☐	NO ☐ (tick appropriate box)
2. As a participant, we understand that:	
a) all essential apparatus, chemicals and other requirements specified in the method must be available in our laboratory when the programme begins;	
b) specified "timing" requirements such as starting date, order of testing specimens and finishing date of the programme must be rigidly met;	
c) the method must be strictly adhered to;	
d) samples must be handled in accordance with instructions;	
e) a qualified operator must perform the measurements.	
Having studied the method and having made a fair appraisal of our capabilities and facilities, we feel that we will be adequately prepared for cooperative testing of this method.	
3. <u>Comments</u>	
(Signed)	
(Company or laboratory)	

Figure 1 — Enlistment questionnaire for interlaboratory study

5.3.3 It should be considered whether it is desirable for some laboratories to obtain some preliminary test results for familiarization with the measurement method before obtaining the official test result and, if so, whether additional material (not precision experiment samples) should be provided for this purpose.

5.3.4 When a material has to be homogenized, this shall be done in the manner most appropriate for that material. When the material to be tested is not homogeneous, it is important to prepare the samples in the manner specified in the method, preferably starting with one batch of commercial material for each level. In the case of unstable materials, special instructions on storage and treatment shall be specified.

5.3.5 For the samples at each level, n separate containers shall be used for each laboratory if there is any danger of the materials deteriorating once the container has been opened (e.g. by oxidation, by losing volatile components, or with hygroscopic material). In the case of unstable materials, special instructions on storage and treatment shall be specified. Precautions may be needed to ensure that samples remain identical up to the time the measurements are made. If the material to be measured consists of a mixture of powders of different relative density or of different grain size, some care is needed because segregation may result from shaking, for example during transport. When reaction with the atmosphere may be expected, the specimens may be sealed into ampoules, either evacuated or filled with an inert gas. For perishable materials such as food or blood samples, it

6 Am Präzisionsversuch beteiligtes Personal

ANMERKUNG 3 Es ist nicht zu erwarten, dass die Arbeitsmethoden in verschiedenen Labors identisch sind. Deshalb soll der Inhalt dieses Abschnitts lediglich als Leitfaden dienen, der in zweckmäßiger Weise modifiziert und auf den einzelnen Fall ausgerichtet werden muss.

6.1 Versuchsgremium

6.1.1 Das Versuchsgremium sollte aus Fachkräften zusammengesetzt sein, die mit dem (vereinheitlichten) Messverfahren und seiner Anwendung vertraut sind.

6.1.2 Die Aufgaben des Versuchsgremiums sind,

- a) den Versuch zu planen und (seinen Ablauf) zu koordinieren;
- b) über die Anzahl der Labors, der Merkmalsniveaus und der auszuführenden Messungen sowie über die Anzahl der zu verlangenden signifikanten Zahlenstellen zu entscheiden;
- c) jemanden für die statistischen Aufgaben zu benennen (siehe 6.2);
- d) jemanden für die Führungsaufgaben zu benennen (siehe 6.3);
- e) die zusätzlich zum vereinheitlichten Messverfahren für die Laborverantwortlichen zu erstellenden Anweisungen zu überlegen;
- f) zu entscheiden, ob einigen Bearbeitern erlaubt werden kann, einige wenige inoffizielle Messungen auszuführen, um nach einer langen Zeitspanne Erfahrung wiederzugewinnen (solche Messungen dürfen niemals an offiziellen Ringversuchs-Proben ausgeführt werden);
- g) den Bericht über die statistische Auswertung zum Zweck der Vervollständigung der Auswertung der Ermittlungsergebnisse zu diskutieren;
- h) Schlusswerte für die Wiederholstandardabweichung und die Vergleichstandardabweichung festzulegen;
- i) zu entscheiden, ob weitere Maßnahmen zur Verbesserung der Vereinheitlichung für das Messverfahren oder im Hinblick auf solche Labors verlangt sind, deren Ergebnisse als Ausreißer rückgewiesen wurden.

6.2 Statistische Aufgaben

Mindestens ein Mitglied des Versuchsgremiums sollte Erfahrung in statistischer Versuchsplanung und in der Auswertung von Versuchen haben. Seine/ihre Aufgaben sind,

- a) sein/ihr Spezialwissen bei der Versuchsplanung einzubringen;
- b) die Daten auszuwerten;
- c) einen Bericht zur Vorlage beim Versuchsgremium zu schreiben, und zwar nach den in 7.7 enthaltenen Anweisungen.

6.3 Führungsaufgaben

6.3.1 Das Organisieren des Versuchs sollte einem einzelnen Labor anvertraut werden. Ein Mitglied des Mitarbeiterstabs dieses Labors sollte die volle Verantwortung dafür übernehmen; er bzw. sie wird „leitende Führungskraft“ genannt und ist bestellt vom Versuchsgremium.

6.3.2 Die Aufgaben der leitenden Führungskraft sind,

- a) die Mitarbeit der benötigten Anzahl von Labors anzuwerben und sicherzustellen, dass (dort) die betreffenden Laborverantwortlichen bestellt sind;
- b) die Vorbereitung der Materialien und Proben zu organisieren und zu überwachen, ebenso die Versendung der Proben; für jedes Merkmalsniveau sollte eine angemessene Menge Material als Reservevorrat zur Seite gelegt werden;
- c) Anleitungen zur Abdeckung aller Gesichtspunkte nach 5.1.2 a) bis 5.1.2 h) zu entwerfen und diese an die Laborverantwortlichen zu verteilen, frühzeitig genug für diese und im Voraus, um Kommentare zu geben und Fragen zu stellen und um sicherzustellen, dass die ausgewählten Bearbeiter solche sind, die üblicherweise bei ihrer Routinearbeit solche Messungen ausführen;
- d) geeignete Formblätter zu entwerfen zur Benutzung als Arbeitsbericht für die Bearbeiter und zur Benutzung für die Laborverantwortlichen zum Bericht über die Ermittlungsergebnisse mit der verlangten Anzahl von signifikanten Zahlenstellen (solche Formblätter können den Namen des Bearbeiters, das Empfangs- und Messdatum der Proben, das benutzte Messgerät und andere relevante Informationen enthalten);
- e) irgendwelche Fragen von Labors im Hinblick auf die Ausführung der Messungen zu behandeln;
- f) darauf zu achten, dass ein Gesamtzeitplan eingehalten wird;
- g) die Formblätter mit den Daten einzusammeln und diese an die statistische Fachkraft weiterzugeben.

6.4 Laborverantwortliche

6.4.1 Ein Mitglied des Mitarbeiterstabes in jedem der teilnehmenden Labors sollte für das Organisieren der eigentlichen Durchführung der Messungen in Übereinstimmung mit den von der leitenden Führungskraft erhaltenen Anweisungen sowie für das Berichten über die Ermittlungsergebnisse verantwortlich gemacht werden.

may be necessary to send them in a deep-frozen state to the participating laboratories with detailed instructions for the procedure for thawing.

6 Personnel involved in a precision experiment

NOTE 3 The methods of operation within different laboratories are not expected to be identical. Therefore the contents of this clause are only intended as a guide to be modified as appropriate to cater for a particular situation.

6.1 Panel

6.1.1 The panel should consist of experts familiar with the measurement method and its application.

6.1.2 The tasks of the panel are:

- a) to plan and coordinate the experiment;
- b) to decide on the number of laboratories, levels and measurements to be made, and the number of significant figures to be required;
- c) to appoint someone for the statistical functions (see 6.2);
- d) to appoint someone for the executive functions (see 6.3);
- e) to consider the instructions to be issued to the laboratory supervisors in addition to the standard measurement method;
- f) to decide whether some operators may be allowed to carry out a few unofficial measurements in order to regain experience of the method after a long interval (such measurements shall never be carried out on the official collaborative samples);
- g) to discuss the report of the statistical analysis on completion of the analysis of the test results;
- h) to establish final values for the repeatability standard deviation and the reproducibility standard deviation;
- i) to decide if further actions are required to improve the standard for the measurement method or with regard to laboratories whose test results have been rejected as outliers.

6.2 Statistical functions

At least one member of the panel should have experience in statistical design and analysis of experiments. His/her tasks are:

- a) to contribute his/her specialized knowledge in designing the experiment;
- b) to analyse the data;
- c) to write a report for submission to the panel following the instructions contained in 7.7.

6.3 Executive functions

6.3.1 The actual organization of the experiment should be entrusted to a single laboratory. A member of the staff of that laboratory should take full responsibility; he/she is called the executive officer and is appointed by the panel.

6.3.2 The tasks of the executive officer are:

- a) to enlist the cooperation of the requisite number of laboratories and to ensure that supervisors are appointed;
- b) to organize and supervise the preparation of the materials and samples and the dispatch of the samples; for each level, an adequate quantity of material should be set aside as a reserve stock;
- c) to draft instructions covering all the points in 5.1.2 a) to h), and circulate them to the supervisors early enough in advance for them to raise any comments or queries and to ensure that operators selected are those who would normally carry out such measurements in routine operations;
- d) to design suitable forms for the operator to use as a working record and for the supervisor to report the test results to the requisite number of significant figures (such forms may include the name of the operator, the dates on which samples were received and measured, the equipment used and any other relevant information);
- e) to deal with any queries from laboratories regarding the performance of the measurements;
- f) to see that an overall time schedule is maintained;
- g) to collect the data forms and present them to the statistical expert.

6.4 Supervisors

6.4.1 A staff member in each of the participating laboratories should be made responsible for organizing the actual performance of the measurements, in keeping with instructions received from the executive officer, and for reporting the test results.

6.4.2 Die Aufgaben des Laborverantwortlichen sind,

- a) sicherzustellen, dass die ausgewählten Bearbeiter solche sind, die üblicherweise bei ihrer Routinearbeit solche Messungen auszuführen haben;
- b) das Aushändigen der Proben an den (die) Bearbeiter in Übereinstimmung mit den Anweisungen der leitenden Führungskraft (und nötigenfalls das Bereitstellen von Material für Versuche zur Einarbeitung);
- c) das Überwachen der Ausführung der Messungen (der Laborverantwortliche muss nicht an der Ausführung der Messungen teilnehmen);
- d) sicherzustellen, dass die Bearbeiter die geforderte Anzahl von Messungen ausführen;
- e) die Befolgung des vorgegebenen Terminplans für die Ausführung der Messungen sicherzustellen;
- f) die Ermittlungsergebnisse zu sammeln, aufgezeichnet entsprechend der vereinbarten Anzahl von Dezimalstellen, eingeschlossen irgendwelche in Erfahrung gebrachte Unregelmäßigkeiten und Schwierigkeiten sowie Kommentare durch die Bearbeiter.

6.4.3 Der Verantwortliche jedes Labors sollte einen vollständigen Bericht schreiben, der die folgende Information enthalten sollte:

- a) die Ermittlungsergebnisse, lesbar eingetragen durch ihren Erzeuger in die bereitgestellten Formblätter, nicht übertragen oder maschinengeschrieben (als Alternative kann der Ausdruck eines Computers oder einer Prüfmaschine annehmbar sein);
- b) die original beobachteten Werte oder Ablesungen (sofern vorhanden), aus denen die Ermittlungsergebnisse abgeleitet sind, lesbar durch den Bearbeiter in die bereitgestellten Formblätter eingetragen, nicht übertragen oder maschinengeschrieben;
- c) Kommentare durch die Bearbeiter über die Vereinheitlichung des Messverfahrens;
- d) Information über Unregelmäßigkeiten oder Störungen, die während der Messungen vorgekommen sein mögen, eingeschlossen irgendwelche Änderungen von Bearbeitern; und zwar zusammen mit einer Darlegung, welche Messungen durch welche Bearbeiter ausgeführt wurden; und schließlich Gründe für irgendwelche fehlenden Ergebnisse;
- e) der Zeitpunkt (die Zeitpunkte), zu dem (zu denen) die Proben in Empfang genommen wurden;
- f) der Zeitpunkt (die Zeitpunkte), zu dem (zu denen) jede Probe gemessen wurde;
- g) Information über die Geräteausrüstung, sofern wesentlich;
- h) irgendeine andere sachdienliche Information.

6.5 Bearbeiter

6.5.1 In jedem Labor müssen die Messungen durch nur einen Mitarbeiter ausgeführt werden, der unter jenen als repräsentativ ausgewählt wurde, die in ihrer üblichen Arbeit geeignet sind, die Messungen auszuführen.

6.5.2 Weil das Ziel des Versuchs die Ermittlung der Präzision ist, wie sie aufgrund der allgemeinen Grundgesamtheit von Bearbeitern gewonnen werden kann, die mit dem vereinheitlichten Messverfahren arbeiten, sollten den Bearbeitern im Allgemeinen keine ergänzenden Erläuterungen zu den Forderungen an das Messverfahren gegeben werden. Allenfalls sollte gegenüber den Bearbeitern zum Ausdruck gebracht werden, dass es Zweck der Übung ist, das Ausmaß herauszufinden, um das sich in der Praxis Ergebnisse unterscheiden können, so dass für sie eine kleinere Versuchung besteht, Ergebnisse auszuondern oder nachzuarbeiten, von denen sie das Gefühl haben, sie seien widersprüchlich.

6.5.3 Obwohl die Bearbeiter üblicherweise keine ergänzenden weiteren Erläuterungen zum vereinheitlichten Messverfahren erhalten sollten, sollten sie ermutigt werden, die Vereinheitlichungsregeln zu kommentieren, insbesondere aber darzulegen, ob die in diesen enthaltenen Anweisungen ausreichend eindeutig und klar sind.

6.5.4 Die Aufgaben der Bearbeiter sind:

- a) Ausführen der Messungen entsprechend dem vereinheitlichten Messverfahren;
- b) Berichten über irgendwelche erlebten Besonderheiten oder Schwierigkeiten; es ist besser, über ein Missgeschick zu berichten, als das Ermittlungsergebnis zurechtzurücken, weil ein oder zwei fehlende Ermittlungsergebnisse den Versuch nicht ungültig machen werden, und weil viele davon eine Unzulänglichkeit des Verfahrens anzeigen;
- c) Kommentare abgeben über die Angemessenheit der Anweisungen in den Vereinheitlichungsregeln; Bearbeiter sollten über alle Fälle berichten, in denen sie nicht in der Lage sind, ihre Anweisungen zu befolgen, zumal dies ebenfalls eine Unzulänglichkeit in den Vereinheitlichungsregeln anzeigen kann.

7 Statistische Auswertung des Präzisionsversuchs

7.1 Einleitende Überlegungen

7.1.1 Die Auswertung der Daten, die als eine statistische Aufgabe betrachtet werden sollte, die durch eine statistische Fachkraft gelöst werden muss, umfasst drei aufeinander folgende Abschnitte:

- a) Kritisches Prüfen der Daten mit dem Ziel, Ausreißer oder andere Unregelmäßigkeiten zu erkennen und zu behandeln, sowie Prüfen der Eignung des (statistischen) Modells;
- b) Errechnen vorläufiger Werte für Präzision und Mittelwerte, getrennt für jedes Merkmalsniveau;
- c) Feststellen der Schlusswerte für Präzision und Mittelwerte, eingeschlossen die Ermittlung einer Beziehung zwischen Präzision und Merkmalsniveau m , sofern die Auswertung anzeigt, dass eine solche Beziehung bestehen kann.

6.4.2 The tasks of the supervisor are:

- a) ensure that the operators selected are those who would normally carry out such measurements in routine operations;
- b) to hand out the samples to the operator(s) in keeping with the instructions of the executive officer (and to provide material for familiarization experiments, if necessary);
- c) to supervise the execution of the measurements (the supervisor shall not take part in performing the measurements);
- d) to ensure that the operators carry out the required number of measurements;
- e) to ensure adherence to the set timetable for performing the measurements;
- f) to collect the test results recorded to the agreed number of decimal places, including any anomalies and difficulties experienced, and comments made by the operators.

6.4.3 The supervisor of each laboratory should write a full report which should contain the following information:

- a) the test results, entered legibly by their originator on the forms provided, not transcribed or typed (computer or testing machine printout may be acceptable as an alternative);
- b) the original observed values or readings (if any) from which the test results were derived, entered legibly by the operator on the forms provided, not transcribed or typed;
- c) comments by the operators on the standard for the measurement method;
- d) information about irregularities or disturbances that may have occurred during the measurements, including any change of operator that may have occurred, together with a statement as to which measurements were performed by which operator, and the reasons for any missing results;
- e) the date(s) on which the samples were received;
- f) the date(s) on which each sample was measured;
- g) information about the equipment used, if relevant;
- h) any other relevant information.

6.5 Operators

6.5.1 In each laboratory the measurements shall be carried out by one operator selected as being representative of those likely to perform the measurements in normal operations.

6.5.2 Because the object of the experiment is to determine the precision obtainable by the general population of operators working from the standard measurement method, in general the operators should not be given amplifications to the standard for the measurement method. However, it should be pointed out to the operators that the purpose of the exercise is to discover the extent to which results can vary in practice, so that there will be less temptation for them to discard or rework results that they feel are inconsistent.

6.5.3 Although normally the operators should receive no supplementary amplifications to the standard measurement method, they should be encouraged to comment on the standard and, in particular, to state whether the instructions contained in it are sufficiently unambiguous and clear.

6.5.4 The tasks of the operators are:

- a) to perform the measurements according to the standard measurement method;
- b) to report any anomalies or difficulties experienced; it is better to report a mistake than to adjust the test results because one or two missing test results will not spoil the experiment and many indicate a deficiency in the standard;
- c) to comment on the adequacy of the instructions in the standard; operators should report any occasions when they are unable to follow their instructions as this may also indicate a deficiency in the standard.

7 Statistical analysis of a precision experiment

7.1 Preliminary considerations

7.1.1 The analysis of the data, which should be considered as a statistical problem to be solved by a statistical expert, involves three successive stages:

- a) critical examination of the data in order to identify and treat outliers or other irregularities and to test the suitability of the model;
- b) computation of preliminary values of precision and means for each level separately;
- c) establishment of final values of precision and means, including the establishment of a relationship between precision and the level m when the analysis indicates that such a relationship may exist.

7.1.2 Bei der Auswertung werden als Erstes, getrennt für jedes Merkmalsniveau, Schätzwerte errechnet für

- die Wiederholvarianz s_r^2
- die Varianz zwischen den Labors s_L^2
- die Vergleichvarianz $s_R^2 = s_r^2 + s_L^2$
- den arithmetischen Mittelwert m .

7.1.3 Die Auswertung schließt eine systematische Anwendung statistischer Ausreißertests ein, für die von der Literatur her eine große Vielfalt verfügbar ist und für die Zwecke dieses Teils von ISO 5725 angewendet werden könnte. Aus praktischen Gründen wurde nur eine begrenzte Anzahl dieser Tests (in diese Norm) aufgenommen. Sie sind in 7.3 erklärt.

7.2 Tabellarische Zusammenstellung der Ergebnisse sowie angewendete Darstellungsweise

7.2.1 Zellen (in einem Ringversuch)

Jede Kombination eines Labors und eines Merkmalsniveaus wird „Zelle des Ringversuchs“ genannt. Im Idealfall bestehen die Ergebnisse eines Ringversuchs mit p Labors und q Merkmalsniveaus aus einer Tabelle mit p mal q Zellen. Jede Zelle enthält n Ergebnisse der Wiederholprüfung. Sie alle können für die Errechnung der Wiederholstandardabweichung und der Vergleichstandardabweichung eingesetzt werden. Allerdings wird in der Praxis dieser Idealfall nicht immer erreicht. Abweichungen entstehen infolge von überflüssigen Daten, fehlenden Daten und Ausreißern.

7.2.2 Überflüssige Daten

Zuweilen kommt es vor, dass ein Labor mehr als die offiziell verlangten n Ermittlungsergebnisse bereitstellt und darüber berichtet. In diesem Fall muss der Laborverantwortliche darüber berichten, warum das geschehen ist, und welches die richtigen Ermittlungsergebnisse sind. Lautet die Antwort, dass alle Ergebnisse gleichermaßen gültig sind, muss aus den verfügbaren Ermittlungsergebnissen eine Zufallsauswahl vorgenommen werden, um die für die Auswertung geplante Anzahl von Ermittlungsergebnissen auszuwählen.

7.2.3 Fehlende Daten

In anderen Fällen können einige Ermittlungsergebnisse fehlen, beispielsweise infolge des Verlusts einer Probe oder infolge eines Fehlers bei der Ausführung der Messung. Die in 7.1 empfohlene Auswertung ist dergestalt, dass vollständig leere Zellen einfach übergangen werden können, während teilweise leere Zellen in das vereinheitlichte Berechnungsverfahren einbezogen werden können.

7.2.4 Ausreißer

Diese sind Eintragungen unter den Originalermittlungsergebnissen oder in den aus diesen entwickelten Tabellen, die so sehr von vergleichbaren Eintragungen in derselben Tabelle abweichen, dass sie als unvereinbar mit den anderen Daten betrachtet werden. Erfahrung hat gelehrt, dass Ausreißer nicht immer vermieden werden können, und dass sie in ähnlicher Art und Weise in Betracht gezo-

gen werden müssen wie die Behandlung von fehlenden Daten.

7.2.5 Ausreißerlabors

Wenn einige unerklärte, stark abweichende Ermittlungsergebnisse auf unterschiedlichen Merkmalsniveaus innerhalb desselben Labors vorkommen, dann kann dieses Labor als Ausreißerlabor betrachtet werden, das eine zu große Varianz innerhalb des Labors und/oder eine zu große systematische Abweichungskomponente im Niveau seiner Ermittlungsergebnisse hat. Es kann dann vernünftig sein, einige oder alle Daten eines solchen Ausreißerlabors auszusondern.

Dieser Teil von ISO 5725 stellt keinen statistischen Test bereit, durch den verdächtige Labors beurteilt werden könnten. Die primäre Entscheidung sollte der Verantwortung der statistischen Fachkraft unterliegen, jedoch sollte im Hinblick auf weitere Maßnahmen über alle rückgewiesenen Labors dem Versuchsgremium berichtet werden.

7.2.6 Falsche Daten

Über offensichtlich falsche Daten sollten Ermittlungen angestellt und sie sollten (gemäß dem Ermittlungsergebnis) korrigiert oder ausgeschlossen werden.

7.2.7 Ergebnisse eines ausgewogenen Versuchs auf einheitlichem Merkmalsniveau

Idealfall sind p Labors, bezeichnet mit dem Index i ($i = 1, 2, \dots, p$), wobei jedes Labor q Merkmalsniveaus untersucht, bezeichnet mit dem Index j ($j = 1, 2, \dots, q$), und wobei auf jedem Merkmalsniveau (jeder ij -Kombination) n Wiederholmessungen ($k = 1, 2, \dots, n$) ausgeführt werden, was zu insgesamt pqn Ermittlungsergebnissen führt. Wegen fehlender (siehe 7.2.3) oder (als Ausreißer von allen anderen Ermittlungsergebnissen) abweichender (siehe 7.2.4) Ermittlungsergebnisse oder wegen Ausreißerlabors (siehe 7.2.5) oder wegen falscher Daten (siehe 7.2.6) wird dieser Idealfall nicht immer erreicht. Unter diesen Umständen sind für abweichende Anzahlen von Ermittlungsergebnissen die in 7.2.8 bis 7.2.10 vorgegebenen Bezeichnungen sowie die Verfahren nach 7.4 zugelassen. Muster für empfehlenswerte Formblätter für die statistische Auswertung sind in Bild 2 angegeben. Aus praktischen Gründen werden sie einfach als Formblätter A, B und C (siehe Bild 2) bezeichnet.

7.2.8 Ursprüngliche Ermittlungsergebnisse

Siehe Formblatt A des Bildes 2, dabei ist

- n_{ij} die Anzahl von Ermittlungsergebnissen in der Zelle für das Labor i auf dem Merkmalsniveau j ;
- y_{ijk} eines dieser Ermittlungsergebnisse ($k = 1, 2, \dots, n_{ij}$);
- p_j die Anzahl der Labors mit Bericht über mindestens ein Ermittlungsergebnis für das Merkmalsniveau j (nach Aussondern von Ermittlungsergebnissen, die als Ausreißer oder als falsch gekennzeichnet wurden).

7.1.2 The analysis first computes, for each level separately, estimates of

- the repeatability variance s_r^2
- the between-laboratory variance s_L^2
- the reproducibility variance $s_R^2 = s_r^2 + s_L^2$
- the mean m .

7.1.3 The analysis includes a systematic application of statistical tests for outliers, a great variety of which are available from the literature and which could be used for the purposes of this part of ISO 5725. For practical reasons, only a limited number of these tests, as explained in 7.3, have been incorporated.

7.2 Tabulation of the results and notation used

7.2.1 Cells

Each combination of a laboratory and a level is called a cell of the precision experiment. In the ideal case, the results of an experiment with p laboratories and q levels consist of a table with pq cells, each containing n replicate test results that can all be used for computing the repeatability standard deviation and the reproducibility standard deviation. This ideal situation is not, however, always attained in practice. Departures occur owing to redundant data, missing data and outliers.

7.2.2 Redundant data

Sometimes a laboratory may carry out and report more than the n test results officially specified. In that case, the supervisor shall report why this was done and which are the correct test results. If the answer is that they are all equally valid, then a random selection should be made from those available test results to choose the planned number of test results for analysis.

7.2.3 Missing data

In other cases, some of the test results may be missing, for example because of loss of a sample or a mistake in performing the measurement. The analysis recommended in 7.1 is such that completely empty cells can simply be ignored, while partly empty cells can be taken into account by the standard computational procedure.

7.2.4 Outliers

These are entries among the original test results, or in the tables derived from them, that deviate so much

from the comparable entries in the same table that they are considered irreconcilable with the other data. Experience has taught that outliers cannot always be avoided and they have to be taken into consideration in a similar way to the treatment of missing data.

7.2.5 Outlying laboratories

When several unexplained abnormal test results occur at different levels within the same laboratory, then that laboratory may be considered to be an outlier, having too high a within-laboratory variance and/or too large a systematic error in the level of its test results. It may then be reasonable to discard some or all of the data from such an outlying laboratory.

This part of ISO 5725 does not provide a statistical test by which suspected laboratories may be judged. The primary decision should be the responsibility of the statistical expert; but all rejected laboratories shall be reported to the panel for further action.

7.2.6 Erroneous data

Obviously erroneous data should be investigated and corrected or discarded.

7.2.7 Balanced uniform-level test results

The ideal case is p laboratories called i ($i = 1, 2, \dots, p$), each testing q levels called j ($j = 1, 2, \dots, q$) with n replicates at each level (each ij combination), giving a total of pqn test results. Because of missing (7.2.3) or deviating (7.2.4) test results, or outlying laboratories (7.2.5) or erroneous data (7.2.6), this ideal situation is not always attained. Under these conditions the notations given in 7.2.8 to 7.2.10 and the procedures of 7.4 allow for differing numbers of test results. Specimens of recommended forms for the statistical analysis are given in figure 2. For convenience, they will be referred to simply as forms A, B and C (of figure 2).

7.2.8 Original test results

See form A of figure 2, where

- n_{ij} is the number of test results in the cell for laboratory i at level j ;
- y_{ijk} is any one of these test results ($k = 1, 2, \dots, n_{ij}$);
- p_j is the number of laboratories reporting at least one test result for level j (after eliminating any test results designated as outliers or as erroneous).

Formblatt A: Empfohlenes Formblatt für die Zusammenstellung der Ursprungsdaten									
Labor	M e r k m a l s n i v e a u								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
..					..				
..					..				
<i>i</i>					<i>y_{ijk}</i>				
..					..				
..									
<i>p</i>									

Formblatt B: Empfohlenes Formblatt für die Zusammenstellung der arithmetischen Mittelwerte									
Labor	M e r k m a l s n i v e a u								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
<i>i</i>					$\bar{y}_{ij}$				
..									
<i>p</i>									

Formblatt C: Empfohlenes Formblatt für die Zusammenstellung der Maße für die Streuung innerhalb einer Zelle									
Labor	M e r k m a l s n i v e a u								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
<i>i</i>					<i>s_{ij}</i>				
..									
<i>p</i>									

Bild 2 — Empfohlene Formblätter für die Zusammenstellung von Ermittlungsergebnissen für die Auswertung

7.2.9 Arithmetische Mittelwerte einer Zelle (Formblatt B des Bildes 2)

Diese ergeben sich aus dem Formblatt A wie folgt:

$$\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \quad \dots (2)$$

Die arithmetischen Mittelwerte einer Zelle sollten mit einer signifikanten Stelle mehr aufgezeichnet sein als die Ermittlungsergebnisse im Formblatt A.

7.2.10 Maße für die Streuung innerhalb einer Zelle (Formblatt C des Bildes 2)

Diese ergeben sich aus dem Formblatt A (siehe 7.2.8) und dem Formblatt B (siehe 7.2.9) wie folgt.

Form A — Recommended form for the collation of the original data									
Laboratory	Level								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
..					..				
..					..				
<i>i</i>					<i>y_{ijk}</i>				
..					..				
..									
<i>p</i>									

Form B — Recommended form for the collation of the means									
Laboratory	Level								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
<i>i</i>					<i>y_{ij}</i>				
..									
<i>p</i>									

Form C — Recommended form for the collation of the measures of spread within cells									
Laboratory	Level								
	1	2	..	..	<i>j</i>	..	..	<i>q</i> - 1	<i>q</i>
1									
2									
..									
<i>i</i>					<i>s_{ij}</i>				
..									
<i>p</i>									

Figure 2 — Recommended forms for the collation of results for analysis

7.2.9 Cell means (form B of figure 2)

These are derived from form A as follows:

$$\bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk} \quad \dots (2)$$

The cell means should be recorded to one more significant figure than the test result in form A.

7.2.10 Measures of cell spread (form C of figure 2)

These are derived from form A (see 7.2.8) and form B (see 7.2.9) as follows.

Für den allgemeinen Fall ist die Standardabweichung innerhalb einer Zelle zu benutzen

$$s_{ij} = \sqrt{\frac{1}{n_{ij}-1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2} \quad \dots (3)$$

oder nach der äquivalenten Formel

$$s_{ij} = \sqrt{\frac{1}{n_{ij}-1} \left[\sum_{k=1}^{n_{ij}} (y_{ijk})^2 - \frac{1}{n_{ij}} \left[\sum_{k=1}^{n_{ij}} y_{ijk} \right]^2 \right]} \quad \dots (4)$$

Bei der Anwendung dieser Formeln muss darauf geachtet werden, eine hinreichende Anzahl Stellen für die Berechnung vorzusehen, das heißt: Jeder Zwischenwert muss mit mindestens doppelt so vielen oder mehr Stellen gerechnet werden, wie die Ursprungsdaten aufweisen.

ANMERKUNG 4 Wenn eine Zelle ij zwei Ermittlungsergebnisse enthält, ist die Standardabweichung innerhalb der Zelle

$$s_{ij} = |y_{ij1} - y_{ij2}| / \sqrt{2} \quad \dots (5)$$

Zur Vereinfachung können also die Beträge der Differenzen anstelle der Standardabweichungen angewendet werden, sofern alle Zellen zwei Ermittlungsergebnisse enthalten.

Die Standardabweichung sollte mit einer signifikanten Stelle mehr ausgedrückt werden als die Ergebnisse im Formblatt A.

Für Werte von n_{ij} unter 2 sollte im Formblatt C ein Strich eingetragen werden.

7.2.11 Korrigierte oder rückgewiesene Daten

Weil einige Daten aufgrund der in 7.1.3, 7.3.3 und 7.3.4 erwähnten Tests korrigiert oder rückgewiesen werden können, können die Werte von y_{ijk} , n_{ij} und p_j , wie sie für die abschließenden Ermittlungen von Präzision und Mittelwert angewendet werden, sich von den Werten unterscheiden, wie sie in den Formblättern A, B und C des Bildes 2 aufgezeichnet sind. Aus diesem Grund muss beim Bericht über die Schlusswerte für Präzision und Richtigkeit gegebenenfalls stets erwähnt werden, welche Daten korrigiert oder ausgesondert worden sind.

7.3 Ergebnisprüfung bezüglich Vereinbarkeit und Ausreißer

Siehe den bibliografischen Hinweis [3].

Wiederhol- und Vergleichstandardabweichungen sind zu schätzen aus den für eine Anzahl spezifischer Merkmalsniveaus zusammengestellten Daten. Die Schätzwerte können sich ändern, weil einzelne Labors oder Werte sich als unvereinbar mit allen anderen Labors oder allen anderen Werten erweisen können. Deshalb müssen Entscheidungen bezüglich (dieser Labors und) dieser Werte getroffen werden. Zwei Ansätze sind eingeführt:

- grafische Vereinbarkeitsprüfungen;
- numerische Ausreißertests.

7.3.1 Grafische Vereinbarkeitsprüfungen

Zwei Prüfgrößen werden angewendet: Sie heißen „Mandels h -Statistik“ und „Mandels k -Statistik“. Es kann angemerkt werden, dass diese sowohl für die Laborbeurteilung hilfreich sind als auch die Variabilität des (vereinheitlichten) Messverfahrens beschreiben.

7.3.1.1 Die Vereinbarkeits-Prüfgröße zwischen den Labors, h , ist für jedes Labor wie folgt zu berechnen: Die Zellenabweichung (arithmetischer Mittelwert einer Zelle minus Gesamtmittelwert für das betreffende Merkmalsniveau) ist durch die Standardabweichung der arithmetischen Mittelwerte einer Zelle (für dieses betrachtete Merkmalsniveau) zu dividieren:

$$h_{ij} = \frac{\bar{y}_{ij} - \bar{y}_j}{\sqrt{\frac{1}{(p_j-1)} \sum_{i=1}^{p_j} (\bar{y}_{ij} - \bar{y}_j)^2}} \quad \dots (6)$$

wobei für $\bar{y}_{ij}$ 7.2.9 und für $\bar{y}_j$ 7.4.4 gilt.

Die Werte h_{ij} für jede Zelle sind abhängig vom Labor aufzuzeichnen, in Gruppen für jedes Merkmalsniveau (und getrennt gruppiert für die unterschiedlichen, durch jedes Labor geprüften Merkmalsniveaus) (siehe Bild B.7).

7.3.1.2 Die Vereinbarkeits-Prüfgröße innerhalb des Labors, k , ist zu berechnen, indem zuerst die „kombinierte Standardabweichung innerhalb der Zellen“ für jedes Merkmalsniveau errechnet wird, also

$$\sqrt{\frac{\sum s_{ij}^2}{p_j}}$$

Dann ist für jedes Labor und für jedes Merkmalsniveau zu berechnen:

$$k_{ij} = \frac{s_{ij} \sqrt{p_j}}{\sqrt{\sum s_{ij}^2}} \quad \dots (7)$$

Die Werte k_{ij} für jede Zelle sind abhängig vom Labor aufzuzeichnen, in Gruppen für jedes Merkmalsniveau (und getrennt gruppiert für die unterschiedlichen, durch jedes Labor geprüften Merkmalsniveaus) (siehe Bild B.8).

For the general case, use the intracell standard deviation

$$s_{ij} = \sqrt{\frac{1}{n_{ij} - 1} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_{ij})^2} \quad \dots (3)$$

or, equivalently

$$s_{ij} = \sqrt{\frac{1}{n_{ij} - 1} \left[\sum_{k=1}^{n_{ij}} (y_{ijk})^2 - \frac{1}{n_{ij}} \left[\sum_{k=1}^{n_{ij}} y_{ijk} \right]^2 \right]} \quad \dots (4)$$

In using these equations, care shall be taken to retain a sufficient number of digits in the calculations; i.e. every intermediate value shall be calculated to at least twice as many digits as in the original data.

NOTE 4 If a cell ij contains two test results, the intracell standard deviation is

$$s_{ij} = |y_{ij1} - y_{ij2}| / \sqrt{2} \quad \dots (5)$$

Therefore, for simplicity, absolute differences can be used instead of standard deviations if all cells contain two test results.

The standard deviation should be expressed to one more significant figure than the results in form A.

For values of n_{ij} less than 2, a dash should be inserted in form C.

7.2.11 Corrected or rejected data

As some of the data may be corrected or rejected on the basis of the tests mentioned in 7.1.3, 7.3.3 and 7.3.4, the values of y_{ijk} , n_{ij} and p_j used for the final determinations of precision and mean may be different from the values referring to the original test results as recorded in forms A, B and C of figure 2. Hence in reporting the final values for precision and trueness, it shall always be stated what data, if any, have been corrected or discarded.

7.3 Scrutiny of results for consistency and outliers

See reference [3].

From data collected on a number of specific levels, repeatability and reproducibility standard deviations are to be estimated. The presence of individual laboratories or values that appear to be inconsistent with all other laboratories or values may change the estimates, and decisions have to be made with respect to these values. Two approaches are introduced:

a) graphical consistency technique;
b) numerical outlier tests.

7.3.1 Graphical consistency technique

Two measures called Mandel's h and k statistics are used. It may be noted that, as well as describing the variability of the measurement method, these help in laboratory evaluation.

7.3.1.1 Calculate the between-laboratory consistency statistic, h , for each laboratory by dividing the cell deviation (cell mean minus the grand mean for that level) by the standard deviation among cell means (for that level):

$$h_{ij} = \frac{\bar{y}_{ij} - \bar{\bar{y}}_j}{\sqrt{\frac{1}{(p_j - 1)} \sum_{i=1}^{p_j} (\bar{y}_{ij} - \bar{\bar{y}}_j)^2}} \quad \dots (6)$$

in which, for $\bar{y}_{ij}$ see 7.2.9, and for $\bar{\bar{y}}_j$ see 7.4.4.

Plot the h_{ij} values for each cell in order of laboratory, in groups for each level (and separately grouped for the several levels examined by each laboratory) (see figure B.7).

7.3.1.2 Calculate the within-laboratory consistency statistic, k , by first calculating the pooled within-cell standard deviation

$$\sqrt{\frac{\sum s_{ij}^2}{p_j}}$$

for each level, and then calculate

$$k_{ij} = \frac{s_{ij} \sqrt{p_j}}{\sqrt{\sum s_{ij}^2}} \quad \dots (7)$$

for each laboratory within each level.

Plot the k_{ij} values for each cell in order of laboratory, in groups for each level (and separately grouped for the several levels examined by each laboratory) (see figure B.8).

7.3.1.3 Die Prüfung der grafischen Darstellungen von h und k kann anzeigen, dass spezielle Labors Ergebnismuster ausweisen, die sich markant von den anderen des Versuchs unterscheiden. Dies zeigt sich durch eine (im Vergleich mit den anderen Labors) durchgängig hohe oder niedrige Streuung innerhalb der Zelle und/oder durch extreme (hohe oder niedrige) arithmetische Mittelwerte einer Zelle über viele Merkmalsniveaus hin. Wenn das vorkommt, sollte Verbindung mit dem speziellen Labor aufgenommen werden, um zu versuchen, die Ursache für das abweichende Erscheinungsbild festzustellen. Abhängig vom betreffenden Ergebnis kann die statistische Fachkraft

- a) die Daten des betreffenden Labors vorläufig zurückhalten;
- b) das Labor bitten, die Messung nochmals auszuführen (falls realisierbar);
- c) die Daten des Labors aus dem Versuch entfernen.

7.3.1.4 In den grafischen Darstellungen der Werte h können unterschiedliche Muster erscheinen. Alle Labors können bei den unterschiedlichen Merkmalsniveaus beides haben, positive oder negative Werte von h . Einzelne Labors können dazu neigen, entweder nur positive oder nur negative Werte h zu liefern, wobei die Anzahl der Labors mit negativen Werten etwa derjenigen entspricht, die positive Werte liefern. Keines dieser Muster ist unüblich oder erfordert Untersuchungen, obwohl das zweite dieser Muster die Vermutung nahe legen kann, dass eine gemeinsame Ursache für die systematische Abweichungskomponente dieser Labors besteht. Wenn andererseits alle Werte h für eines der Labors das eine Vorzeichen haben und die Werte h für die anderen Labors haben alle das andere Vorzeichen, dann sollte die Ursache gesucht werden. Ebenso sollte die Ursache gesucht werden, wenn die Werte h vom Merkmalsniveau des Versuchs in systematischer Weise abzuhängen scheinen. Entsprechend den in 8.3 (in den Tabellen 6 und 7) angegebenen Indikatoren sind Linien in die grafischen Darstellungen der Werte h einzuzichnen. Diese Indikatorlinien dienen als Anhaltspunkte bei der Prüfung der Daten auf Muster.

7.3.1.5 Wenn eines der Labors dadurch unter den k Aufzeichnungen auffällt, dass es viele große Werte hat, dann sollte die Ursache gesucht werden: Dies weist darauf hin, dass es eine schlechtere Wiederholpräzision hat als die anderen Labors. Einem Labor könnte das Anlass sein, zu durchweg kleineren Werten k zu kommen, und zwar aufgrund solcher Faktoren wie übertriebene Rundung seiner Daten oder eine sehr grobe Messskala. Entsprechend den in 8.3 (in den Tabellen 6 und 7) angegebenen Indikatoren sind Linien in die grafischen Darstellungen der Werte k einzuzichnen. Diese Indikatorlinien dienen als Anhaltspunkte bei der Prüfung der Daten auf Muster.

7.3.1.6 Wenn eine nach Labors geordnete Darstellung der Werte von h und k zur Vermutung Anlass gibt, dass eines der Labors einige Werte h oder k in der Nähe der Linie kritischer Werte hat, sollte die betreffende Darstellung untersucht werden, gruppiert nach Merkmalsniveaus. Häufig verwandelt sich ein Wert, der in der nach Labors geordneten Darstellung groß erschien, zu einem mit anderen Labors für dasselbe Merkmalsniveau recht gut übereinstimmenden Wert. Wenn aber dieser Wert als stark von

den Werten der anderen Labors abweichend erkannt ist, dann sollte die Ursache gesucht werden.

7.3.1.7 Zusätzlich zu diesen Darstellungen von h und k können Histogramme der arithmetischen Mittelwerte und der Spannweiten der Zellen das Vorhandensein von beispielsweise zwei unterschiedlichen Grundgesamtheiten erkennen lassen. Ein solcher Fall könnte eine spezielle Behandlung verlangen, weil das allgemein für die hier beschriebene Methode zugrunde liegende Prinzip eine einzige, eingipflige Verteilung der Grundgesamtheit voraussetzt.

7.3.2 Numerische Methode der Ausreißerprüfung

7.3.2.1 Für die Behandlung von Ausreißern wird die folgende Verfahrensweise empfohlen:

- a) Die in 7.3.3 und 7.3.4 empfohlenen Tests werden angewendet, um Fastausreißer oder Ausreißer zu erkennen:
 - sofern die Prüfgröße (einen Wert hat, der) kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, wird die geprüfte Einheit als korrekt angenommen;
 - sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 5 % ist, aber kleiner oder höchstens gleich der Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 %, wird die geprüfte Einheit als „Fastausreißer“ bezeichnet und mit einem Einzelstern gekennzeichnet;
 - sofern die Prüfgröße (einen Wert hat, der) größer als die Grenze des kritischen Bereichs für ein Signifikanzniveau von 1 % ist, wird die geprüfte Einheit als „statistischer Ausreißer“ bezeichnet und mit einem Doppelstern gekennzeichnet.
- b) Als Nächstes ist zu untersuchen, ob die Fastausreißer und/oder die statistischen Ausreißer durch irgendein technisches Versehen erklärt werden können, beispielsweise durch
 - ein Fehlverhalten bei der Ausführung der Messung,
 - einen Rechenfehler,
 - einen einfachen Schreibfehler beim Aufschreiben des Ermittlungsergebnisses oder
 - die Prüfung einer falschen Probe.

War der Fehler ein Rechenfehler oder ein Schreibfehler, sollte das verdächtige Ergebnis durch den korrigierten Wert ersetzt werden. War der Fehler durch die Prüfung einer falschen Probe verursacht, sollte das Ergebnis in seine zutreffende Zelle platziert werden. Ist diese Korrektur ausgeführt worden, sollte die Prüfung auf Fastausreißer oder Ausreißer wiederholt werden. Wenn die Erklärung des technischen Fehlers von der Art ist, dass damit die Unmöglichkeit des Ersetzens des verdächtigen Ermittlungsergebnisses bewiesen ist, dann sollte der Wert als ein „echter“ Ausreißer ausgesondert werden, der eigentlich nicht zum Versuch gehört.

- c) Wenn irgendwelche Fastausreißer und/oder Ausreißer übrig bleiben, die nicht erklärt werden können, sollte ein Ausreißerlabor gehörig rückgewiesen werden.