

DIN ISO 16269-6

ICS 03.120.30

Ersatz für
DIN 55303-5:1987-02**Statistische Auswertung von Daten –
Teil 6: Ermittlung von statistischen Anteilsbereichen (ISO 16269-6:2005);
Text Deutsch und Englisch**

Statistical interpretation of data –
Part 6: Determination of statistical tolerance intervals (ISO 16269-6:2005);
Text in German and English

Interprétation statistique des données –
Partie 6: Détermination des intervalles statistiques de tolérance (ISO 16269-6:2005);
Texte en allemand et anglais

Gesamtumfang 58 Seiten

Normenausschuss Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ) im DIN

Nationales Vorwort

Dieses Dokument wurde vom Technischen Komitee ISO/TC 69, *Applications of statistical methods*, erarbeitet. Der zuständige nationale Normenausschuss ist der NA 147-00-02 AA, *Qualitätsmanagement, Statistik und Zertifizierungsgrundlagen (NQSZ)*, Arbeitsausschuss *Angewandte Statistik*.

Es wird auf die Möglichkeit hingewiesen, dass für Teile dieses Dokuments Patentrechte bestehen. Das DIN übernimmt keine Verantwortung für die Kennzeichnung solcher Rechte.

Diese erste Ausgabe von ISO 16269-6 annulliert und ersetzt ISO 3207:1975, die technisch überarbeitet worden ist.

Die Normenreihe ISO 16269 umfasst unter dem Haupttitel *Statistische Auswertung von Daten* die folgenden Teile, von denen sich einige noch in Erarbeitung befinden:

- *Teil 1: Leitfaden zur statistischen Auswertung von Daten*
- *Teil 2: Darstellung statistischer Daten*
- *Teil 3: Tests auf Abweichung von der Normalverteilung*
- *Teil 4: Erkennung und Behandlung von Ausreißern*
- *Teil 5: Schätzung und Tests für Mittelwerte und Varianzen für die Normalverteilung, mit Gütefunktionen der Tests*
- *Teil 6: Ermittlung von statistischen Anteilsbereichen*
- *Teil 7: Median — Punktschätzung und Vertrauensbereiche*
- *Teil 8: Ermittlung von Prognosebereichen*

Änderungen

Gegenüber DIN 55303-5:1987-02 wurden folgende Änderungen vorgenommen:

- a) die Norm ist vollständig neu gestaltet worden.

Frühere Ausgaben

DIN 55303-5: 1983-03, 1987-02

— Leerseite —

Statistische Auswertung von Daten — Teil 6: Ermittlung von statistischen Anteilsbereichen

Inhalt

	Seite
Nationales Vorwort	2
Einleitung.....	6
1 Anwendungsbereich	10
2 Normative Verweisungen.....	10
3 Begriffe und Formelzeichen	10
3.1 Begriffe	10
3.2 Formelzeichen.....	12
4 Verfahren	14
4.1 Normalverteilte Grundgesamtheit mit bekannter Varianz und bekanntem Erwartungswert	14
4.2 Normalverteilte Grundgesamtheit mit bekannter Varianz und unbekanntem Erwartungswert.....	14
4.3 Normalverteilte Grundgesamtheit mit unbekannter Varianz und unbekanntem Erwartungswert.....	14
4.4 Beliebige kontinuierliche Verteilung unbekannten Typs	14
5 Beispiele	16
5.1 Daten	16
5.2 Beispiel 1: Einseitig begrenzter statistischer Anteilsbereich bei bekannter Varianz.....	16
5.3 Beispiel 2: Zweiseitig begrenzter statistischer Anteilsbereich bei bekannter Varianz	18
5.4 Beispiel 3: Einseitig begrenzter statistischer Anteilsbereich bei unbekannter Varianz	18
5.5 Beispiel 4: Zweiseitig begrenzter statistischer Anteilsbereich bei unbekannter Varianz.....	20
5.6 Beispiel 5: Verteilungsfreier statistischer Anteilsbereich für kontinuierliche Verteilung.....	20
Anhang A (informativ) Formblätter für Anteilsbereiche	24
Anhang B (normativ) Faktoren $k_1(n; p; 1 - \alpha)$ für die Grenzen einseitig begrenzter statistischer Anteilsbereiche bei bekanntem σ	36
Anhang C (normativ) Faktoren $k_2(n; p; 1 - \alpha)$ für die Grenzen zweiseitig begrenzter statistischer Anteilsbereiche bei bekanntem σ	40
Anhang D (normativ) Faktoren $k_3(n; p; 1 - \alpha)$ für die Grenzen einseitig begrenzter statistischer Anteilsbereiche bei unbekanntem σ	44
Anhang E (normativ) Faktoren $k_4(n; p; 1 - \alpha)$ für die Grenzen zweiseitig begrenzter statistischer Anteilsbereiche bei unbekanntem σ	48
Anhang F (normativ) Verteilungsfreie einseitig begrenzte statistische Anteilsbereiche	52
Anhang G (normativ) Verteilungsfreie zweiseitig begrenzte statistische Anteilsbereiche	53
Anhang H (informativ) Herleitung eines verteilungsfreien statistischen Anteilsbereiches bei beliebiger Verteilung	54
Anhang I (informativ) Herleitung von Faktoren für zweiseitig begrenzte parametrische statistische Anteilsbereiche	56
Literaturhinweise	58

Statistical interpretation of data — Part 6: Determination of statistical tolerance intervals

Contents

	Page
Introduction.....	7
1 Scope	11
2 Normative references	11
3 Terms, definitions and symbols	11
3.1 Terms and definitions	11
3.2 Symbols.....	13
4 Procedures	15
4.1 Normal population with known variance and known mean	15
4.2 Normal population with known variance and unknown mean.....	15
4.3 Normal population with unknown variance and unknown mean	15
4.4 Any continuous distribution of unknown type.....	15
5 Examples	17
5.1 Data	17
5.2 Example 1: One-sided statistical tolerance interval under known variance	17
5.3 Example 2: Two-sided statistical tolerance interval under known variance.....	19
5.4 Example 3: One-sided statistical tolerance interval under unknown variance.....	19
5.5 Example 4: Two-sided statistical tolerance interval under unknown variance	21
5.6 Example 5: Distribution-free statistical tolerance interval for continuous distribution	21
Annex A (informative) Forms for tolerance intervals	25
Annex B (normative) One-sided statistical tolerance limit factors, $k_1(n; p; 1 - \alpha)$, for known σ	36
Annex C (normative) Two-sided statistical tolerance limit factors, $k_2(n; p; 1 - \alpha)$, for known σ	40
Annex D (normative) One-sided statistical tolerance limit factors, $k_3(n; p; 1 - \alpha)$, for unknown σ	44
Annex E (normative) Two-sided statistical tolerance limit factors, $k_4(n; p; 1 - \alpha)$, for unknown σ	48
Annex F (normative) One-sided distribution-free statistical tolerance intervals	52
Annex G (normative) Two-sided distribution-free statistical tolerance intervals	53
Annex H (informative) Construction of a distribution-free statistical tolerance interval for any type of distribution	55
Annex I (informative) Computation of factors for two-sided parametric statistical tolerance intervals	57
Bibliography.....	58

Einleitung

Ein statistischer Anteilsbereich ist ein aus einer Stichprobe geschätzter Bereich, von dem gesagt werden kann, dass er mit einer Wahrscheinlichkeit von $1 - \alpha$, beispielsweise 95 %, mindestens einen festgelegten Anteil p von Einheiten der Grundgesamtheit umfasst. Die Grenzen eines statistischen Anteilsbereiches werden statistische Anteilsgrenzen genannt. Das Vertrauensniveau $1 - \alpha$ ist die Wahrscheinlichkeit, dass ein in der beschriebenen Art konstruierter statistischer Anteilsbereich mindestens einen Anteil p der Grundgesamtheit umfassen wird. Umgekehrt ist die Wahrscheinlichkeit gleich α , dass dieser Bereich weniger als den Anteil p der Grundgesamtheit umfassen wird. Dieser Teil von ISO 16269 beschreibt sowohl einseitig begrenzte als auch zweiseitig begrenzte statistische Anteilsbereiche. Ein einseitig begrenzter Anteilsbereich hat eine obere oder eine untere Grenze, während ein zweiseitig begrenzter Anteilsbereich sowohl eine obere als auch eine untere Grenze hat.

Anteilsbereiche sind Funktionen der beobachteten Werte der Stichprobe, das heißt Kenngrößen, und sie werden normalerweise für verschiedene Stichproben unterschiedliche Werte annehmen. Damit die in diesem Teil von ISO 16269 bereitgestellten Verfahren gelten, müssen die beobachteten Werte voneinander unabhängig sein.

In diesem Teil von ISO 16269 werden zwei Arten von Anteilsbereichen vorgestellt, parametrische und verteilungsfreie. Der parametrische Ansatz geht von der Annahme aus, dass das in der Grundgesamtheit untersuchte Merkmal normalverteilt ist. Daher kann für die Wahrscheinlichkeit, dass der berechnete Anteilsbereich mindestens einen Anteil p der Grundgesamtheit umfasst, nur dann $1 - \alpha$ unterstellt werden, wenn die Annahme der Normalverteilung zutrifft. Für normalverteilte Merkmale wird der statistische Anteilsbereich durch Verwendung eines der Formblätter A, B, C oder D in Anhang A ermittelt.

Parametrische Verfahren für andere Verteilungen als die Normalverteilung werden in diesem Teil von ISO 16269-6 nicht betrachtet. Wenn vermutet wird, dass die Grundgesamtheit nicht normalverteilt ist, darf ein verteilungsfreies Verfahren zur Berechnung statistischer Anteilsbereiche verwendet werden. Das Verfahren zur Ermittlung eines statistischen Anteilsbereiches für eine beliebige kontinuierliche Verteilung ist in den Formblättern E und F von Anhang A enthalten.

Die in diesem Teil von ISO 16269 diskutierten Anteilsgrenzen können verwendet werden, um im statistischen Prozessmanagement die natürliche Fähigkeit eines Prozesses mit einer oder zwei vorgegebenen Spezifikationsgrenzen zu vergleichen, entweder mit einer oberen U , einer unteren L oder mit beiden. Das kommt darin zum Ausdruck, dass diese Anteilsgrenzen auch natürliche Prozessgrenzen genannt worden sind. Siehe dazu ISO 3534-2:1993, 3.2.4, und die allgemeinen Bemerkungen in ISO 3207, die zurückgezogen und durch diesen Teil von ISO 16269 ersetzt wird.

Oberhalb der oberen Spezifikationsgrenze U liegt der obere Anteil p_U fehlerhafter Einheiten (ISO 3534-2:—, 3.2.5.5 und 3.3.1.4) und unter der unteren Spezifikationsgrenze L liegt der untere Anteil p_L fehlerhafter Einheiten (ISO 3534-2:—, 3.2.5.6 und 3.3.1.5). Die Summe $p_U + p_L = p_T$ wird Gesamtanteil fehlerhafter Einheiten genannt (ISO 3534-2:—, 3.2.5.7). Zwischen den Spezifikationsgrenzen U und L liegt der Anteil fehlerfreier Einheiten $1 - p_T$.

Im statistischen Prozessmanagement sind die Grenzen U und L vorab festgelegt und die Anteile p_U , p_L und p_T sind entweder berechnet, wenn die Verteilung als bekannt angenommen wird, oder sie werden anderweitig geschätzt. Es gibt viele Anwendungen für statistische Anteilsbereiche, wenn auch die obige ein Beispiel für ein Qualitätslenkungsproblem ist. Umfassendere Anwendungen und weitere statistische Anteilsbereiche werden in vielen Lehrbüchern eingeführt, wie zum Beispiel in Hahn und Meeker [10].

Introduction

A statistical tolerance interval is an estimated interval, based on a sample, which can be asserted with confidence $1 - \alpha$, for example 95 %, to contain at least a specified proportion p of the items in the population. The limits of a statistical tolerance interval are called statistical tolerance limits. The confidence level $1 - \alpha$ is the probability that a statistical tolerance interval constructed in the prescribed manner will contain at least a proportion p of the population. Conversely, the probability that this interval will contain less than the proportion p of the population is α . This part of ISO 16269 describes both one-sided and two-sided statistical tolerance intervals; a one-sided interval is constructed with an upper or a lower limit while a two-sided interval is constructed with both an upper and a lower limit.

Tolerance intervals are functions of the observations of the sample, i.e. statistics, and they will generally take different values for different samples. It is necessary that the observations be independent for the procedures provided in this part of ISO 16269 to be valid.

Two types of tolerance interval are provided in this part of ISO 16269, parametric and distribution-free. The parametric approach is based on the assumption that the characteristic being studied in the population has a normal distribution; hence the confidence that the calculated statistical tolerance interval contains at least a proportion p of the population can only be taken to be $1 - \alpha$ if the normality assumption is true. For normally distributed characteristics, the statistical tolerance interval is determined using one of the Forms A, B, C or D given in Annex A.

Parametric methods for distributions other than the normal are not considered in this part of ISO 16269. If departure from normality is suspected in the population, distribution-free statistical tolerance intervals may be constructed. The procedure for the determination of a statistical tolerance interval for any continuous distribution is provided in Forms E and F of Annex A.

The tolerance limits discussed in this part of ISO 16269 can be used to compare the natural capability of a process with one or two given specification limits, either an upper one U or a lower one L or both in statistical process management. An indication of this is the fact that these tolerance limits have also been called natural process limits. See ISO 3534-2:1993, 3.2.4, and the general remarks in ISO 3207 which will be cancelled and replaced by this part of ISO 16269.

Above the upper specification limit U there is the upper fraction nonconforming p_U (ISO 3534-2:—, 3.2.5.5 and 3.3.1.4) and below the lower specification limit L there is the lower fraction nonconforming p_L (ISO 3534-2:—, 3.2.5.6 and 3.3.1.5). The sum $p_U + p_L = p_T$ is called the total fraction nonconforming. (ISO 3534-2:—, 3.2.5.7). Between the specification limits U and L there is the fraction conforming $1 - p_T$.

In statistical process management the limits U and L are fixed in advance and the fractions p_U , p_L and p_T are either calculated, if the distribution is assumed to be known, or otherwise estimated. There are many applications of statistical tolerance intervals, although the above shows an example to a quality control problem. Wider applications and more statistical intervals are introduced in many textbooks such as Hahn and Meeker ^[10].

Im Gegensatz dazu sind für die in diesem Teil von ISO 16269 betrachteten Anteilsbereiche das Vertrauensniveau für den Intervallschätzer und der Anteil der Verteilung innerhalb des Bereiches (der dem oben erwähnten Anteil fehlerfreier Einheiten entspricht) vorab festgelegt, und die Grenzen werden geschätzt. Diese Grenzen dürfen mit U und L verglichen werden. Daher kann die Eignung der gegebenen Spezifikationsgrenzen U und L mit den vorliegenden Eigenschaften des Prozesses verglichen werden. Die einseitig begrenzten Anteilsbereiche werden verwendet, wenn entweder nur die obere oder nur die untere Spezifikationsgrenze von Bedeutung ist, während die zweiseitig begrenzten Bereiche verwendet werden, wenn sowohl die obere als auch die untere Spezifikationsgrenze gleichzeitig betrachtet werden.

Die Terminologie bezüglich dieser unterschiedlichen Grenzen und Bereiche ist verwirrend gewesen, da die „Spezifikationsgrenzen“ früher auch „Toleranzgrenzen“ genannt wurden (siehe dazu die Terminologienorm 3534-2:1993, 1.4.3, in der diese beiden Benennungen ebenso wie die Benennung „Grenzwerte“ alle als synonym für diesen Begriff verwendet wurden). Bei der letzten Überarbeitung von ISO 3534-2:— ist nur die Benennung Spezifikationsgrenzen für diesen Begriff beibehalten worden. Des Weiteren verwendet der *Leitfaden zur Angabe der Unsicherheit beim Messen* [5] die Benennung „Erweiterungsfaktor“ (coverage factor), der definiert ist als ein „Zahlenfaktor, mit dem die kombinierte Standardunsicherheit multipliziert wird, um eine erweiterte Messunsicherheit zu erhalten“. Diese Verwendung von „coverage“ unterscheidet sich von der Verwendung dieser Benennung in diesem Teil von ISO 16269.

In contrast, for the tolerance intervals considered in this part of ISO 16269, the confidence level for the interval estimator and the proportion of the distribution within the interval (corresponding to the fraction conforming mentioned above) are fixed in advance, and the limits are estimated. These limits may be compared with U and L . Hence the appropriateness of the given specification limits U and L can be compared with the actual properties of the process. The one-sided tolerance intervals are used when only either the upper specification limit U or the lower specification limit L is relevant, while the two-sided intervals are used when both the upper and the lower specification limits are considered simultaneously.

The terminology with regard to these different limits and intervals has been confusing as the “specification limits” were earlier also called “tolerance limits” (see the terminology standard ISO 3534-2:1993, 1.4.3, where both these terms as well as the term “limiting values” were all used as synonyms for this concept). In the latest revision of ISO 3534-2:—, only the term specification limits have been kept for this concept. Furthermore, the *Guide for the expression of uncertainty in measurement* [5] uses the term “coverage factor” defined as a “numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty”. This use of “coverage” differs from the use of the term in this part of ISO 16269.

1 Anwendungsbereich

Dieser Teil von ISO 16269 beschreibt Verfahren für die Berechnung von Anteilsbereichen, die mit einem vorgegebenen Vertrauensniveau mindestens einen festgelegten Anteil der Grundgesamtheit umfassen. Sowohl einseitig begrenzte als auch zweiseitig begrenzte Anteilsbereiche werden behandelt, wobei ein einseitig begrenzter Anteilsbereich entweder eine obere oder eine untere Grenze hat, während ein zweiseitig begrenzter Anteilsbereich sowohl eine obere als auch eine untere Grenze hat. Zwei Verfahren werden angeboten, ein parametrisches für den Fall, dass das untersuchte Merkmal normalverteilt ist, und ein verteilungsfreies für den Fall, dass über die Verteilung nichts bekannt ist, außer, dass sie kontinuierlich ist.

2 Normative Verweisungen

Die folgenden zitierten Dokumente sind für die Anwendung dieses Dokuments erforderlich. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability*^{N1)}

ISO 3534-2:—¹⁾, ^{N2)}, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

3 Begriffe und Formelzeichen

3.1 Begriffe

Für die Anwendung dieses Dokuments gelten die Begriffe nach ISO 3534-1 und ISO 3534-2 sowie die folgenden Begriffe.

3.1.1

statistischer Anteilsbereich

Bereich, der derart aus einer Zufallsstichprobe bestimmt wird, dass er bei vorgegebenem Vertrauensniveau mindestens einen festgelegten Anteil der gesamten Grundgesamtheit überdeckt

ANMERKUNG Das Vertrauensniveau ist in diesem Zusammenhang der Langzeitanteil von in dieser Art und Weise konstruierten Bereichen, die mindestens den vorgegebenen Anteil der betrachteten Grundgesamtheit enthalten.

3.1.2

statistische Anteilsgrenze

Kenngroße, die einen Endpunkt eines statistischen Anteilsbereiches darstellt

ANMERKUNG Statistische Anteilsbereiche können entweder einseitig sein, wobei sie entweder eine obere oder eine untere statistische Anteilsgrenze haben, oder zweiseitig sein, in welchem Fall sie beide Grenzen haben.

3.1.3

Abdeckung

Anteil der Einheiten einer Grundgesamtheit, der in einem statistischen Anteilsbereich liegt

ANMERKUNG Dieser Begriff darf nicht mit dem im Leitfaden zur Angabe der Unsicherheit beim Messen (GUM) [5] verwendeten Begriff Erweiterungsfaktor (coverage factor) verwechselt werden.

1) In Überarbeitung. (Revision von ISO 3534-2:1993)

N1) Nationale Fußnote: Der Untertitel wurde in der englischen Fassung falsch angegeben. Dies ist im deutschen Text korrigiert worden.

N2) Nationale Fußnote: ISO 3534-2 ist mit Ausgabedatum 2006 erschienen.

1 Scope

This part of ISO 16269 describes procedures for establishing tolerance intervals that include at least a specified proportion of the population with a specified confidence level. Both one-sided and two-sided statistical tolerance intervals are provided, a one-sided interval having either an upper or a lower limit while a two-sided interval has both upper and lower limits. Two methods are provided, a parametric method for the case where the characteristic being studied has a normal distribution and a distribution-free method for the case where nothing is known about the distribution except that it is continuous.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

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

ISO 3534-2:—¹⁾, *Statistics — Vocabulary and symbols — Part 2: Applied statistics*

3 Terms, definitions and symbols

3.1 Terms and definitions

For the purposes of this document, the terms and definition given in ISO 3534-1, ISO 3534-2 and the following apply.

3.1.1

statistical tolerance interval

interval determined from a random sample in such a way that one may have a specified level of confidence that the interval covers at least a specified proportion of the sampled population

NOTE The confidence level in this context is the long-run proportion of intervals constructed in this manner that will include at least the specified proportion of the sampled population.

3.1.2

statistical tolerance limit

statistic representing an end point of a statistical tolerance interval

NOTE Statistical tolerance intervals can be either one-sided, in which case they have either an upper or a lower statistical tolerance limit, or two-sided, in which case they have both.

3.1.3

coverage

proportion of items in a population lying within a statistical tolerance interval

NOTE This concept is not to be confused with the concept coverage factor used in the *Guide for the expression of uncertainty in measurement* (GUM) [5].

¹⁾ To be published. (Revision of ISO 3534-2:1993)

3.1.4

normalverteilte Grundgesamtheit

Grundgesamtheit, deren Werte einer Normalverteilung folgen

3.2 Formelzeichen

Für diesen Teil von ISO 16269 gelten die folgenden Formelzeichen.

i	laufender Index einer Beobachtung
$k_1(n; p; 1 - \alpha)$	zur Ermittlung von x_L oder x_U verwendeter Faktor für einen einseitig begrenzten Anteilsbereich, wenn der Wert von σ bekannt ist
$k_2(n; p; 1 - \alpha)$	zur Ermittlung von x_L und x_U verwendeter Faktor für einen zweiseitig begrenzten Anteilsbereich, wenn der Wert von σ bekannt ist
$k_3(n; p; 1 - \alpha)$	zur Ermittlung von x_L oder x_U verwendeter Faktor für einen einseitig begrenzten Anteilsbereich, wenn der Wert von σ unbekannt ist
$k_4(n; p; 1 - \alpha)$	zur Ermittlung von x_L und x_U verwendeter Faktor für einen zweiseitig begrenzten Anteilsbereich, wenn der Wert von σ unbekannt ist
n	Anzahl der Beobachtungen in der Stichprobe
p	kleinster Wert des Anteils der Grundgesamtheit, von dem gesagt wird, dass er im statistischen Anteilsbereich liegt
u_p	p -Quantil der standardisierten Normalverteilung
x_i	i -ter beobachteter Wert ($i = 1, 2, \dots, n$)
$x_{\max}$	Maximum der beobachteten Werte: $x_{\max} = \max \{x_1, x_2, \dots, x_n\}$
$x_{\min}$	Minimum der beobachteten Werte: $x_{\min} = \min \{x_1, x_2, \dots, x_n\}$
x_L	untere Grenze des statistischen Anteilsbereichs
x_U	obere Grenze des statistischen Anteilsbereichs
$\bar{x}$	Stichprobenmittelwert, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
s	Stichprobenstandardabweichung, $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}{n(n-1)}}$
$1 - \alpha$	Vertrauensniveau, das mit dem Anspruch verbunden ist, dass der Anteil der Grundgesamtheit innerhalb des Anteilsbereiches größer oder gleich dem spezifizierten Anteil p ist
μ	Erwartungswert der Grundgesamtheit
σ	Standardabweichung der Grundgesamtheit

3.1.4**normal population**

normally distributed population

3.2 Symbols

For the purposes of this part of ISO 16269, the following symbols apply.

i	suffix of an observation
$k_1(n; p; 1 - \alpha)$	factor used to determine x_L or x_U when the value of σ is known for one-sided tolerance interval
$k_2(n; p; 1 - \alpha)$	factor used to determine x_L and x_U when the value of σ is known for two-sided tolerance interval
$k_3(n; p; 1 - \alpha)$	factor used to determine x_L or x_U when the value of σ is unknown for one-sided tolerance interval
$k_4(n; p; 1 - \alpha)$	factor used to determine x_L and x_U when the value of σ is unknown for two-sided tolerance interval
n	number of observations in the sample
p	minimum proportion of the population claimed to be lying in the statistical tolerance interval
u_p	p -fractile of the standard normal distribution
x_i	i th observed value ($i = 1, 2, \dots, n$)
$x_{\max}$	maximum value of the observed values: $x_{\max} = \max \{x_1, x_2, \dots, x_n\}$
$x_{\min}$	minimum value of the observed values: $x_{\min} = \min \{x_1, x_2, \dots, x_n\}$
x_L	lower limit of the statistical tolerance interval
x_U	upper limit of the statistical tolerance interval
$\bar{x}$	sample mean, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
s	sample standard deviation; $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}{n(n-1)}}$
$1 - \alpha$	confidence level for the claim that the proportion of the population lying within the tolerance interval is greater than or equal to the specified level p
μ	population mean
σ	population standard deviation

4 Verfahren

4.1 Normalverteilte Grundgesamtheit mit bekannter Varianz und bekanntem Erwartungswert

Wenn die Werte des Erwartungswertes μ und der Varianz σ^2 einer normalverteilten Grundgesamtheit bekannt sind, dann ist die Verteilung des untersuchten Merkmals vollständig festgelegt. Genau ein Anteil p der Grundgesamtheit liegt

- a) rechts von $x_L = \mu - u_p \times \sigma$ (einseitig nach unten begrenzter Anteilsbereich);
- b) links von $x_U = \mu + u_p \times \sigma$ (einseitig nach oben begrenzter Anteilsbereich);
- c) zwischen $x_L = \mu - u_{(1+p)/2} \times \sigma$ und $x_U = \mu + u_{(1+p)/2} \times \sigma$ (zweiseitig begrenzter Bereich).

ANMERKUNG Da solche Aussagen bekanntermaßen wahr sind, wird ihnen eine Wahrscheinlichkeit von 100 % zugeordnet.

In den obigen Gleichungen ist u_p das p -Quantil der standardisierten Normalverteilung. Zahlenwerte für u_p können der jeweils letzten Zeile der Tabellen B.1 bis B.6 und C.1 bis C.6 entnommen werden.

4.2 Normalverteilte Grundgesamtheit mit bekannter Varianz und unbekanntem Erwartungswert

Formblätter A und B in Anhang A sind in dem Fall anwendbar, wenn die Varianz der normalverteilten Grundgesamtheit bekannt ist, während der Erwartungswert unbekannt ist. Formblatt A gilt für den einseitig begrenzten Fall, während Formblatt B für den zweiseitig begrenzten Fall gilt.

4.3 Normalverteilte Grundgesamtheit mit unbekannter Varianz und unbekanntem Erwartungswert

Formblätter C und D in Anhang A sind in dem Fall anwendbar, wenn sowohl der Erwartungswert als auch die Varianz der normalverteilten Grundgesamtheit unbekannt sind. Formblatt C gilt für den einseitig begrenzten Fall, während Formblatt D für den zweiseitig begrenzten Fall gilt.

4.4 Beliebige kontinuierliche Verteilung unbekannten Typs

Wenn das betrachtete Merkmal eine kontinuierliche Variable aus einer Grundgesamtheit unbekannter Form ist, und wenn eine Stichprobe von n voneinander unabhängigen zufälligen Beobachtungen des Merkmals gezogen worden ist, dann kann ein statistischer Anteilsbereich aus den der Größe nach geordneten Beobachtungen ermittelt werden. Die in den Formblättern E und F von Anhang A angegebenen Verfahren liefern die Ermittlung der Abdeckung oder des Stichprobenumfangs, die oder der für die durch die Extremwerte $x_{\min}$ oder $x_{\max}$ der beobachteten Stichprobe ermittelten Anteilsbereiche benötigt wird, wenn der Vertrauensbereich $1 - \alpha$ gegeben ist.

ANMERKUNG Statistische Anteilsbereiche, die nicht von der Form der untersuchten Grundgesamtheit abhängen, heißen *verteilungsfreie* Anteilsbereiche.

Dieser Teil von ISO 16269 bietet, außer für die Normalverteilung, keine Verfahren für andere Verteilungen bekannten Typs. Wenn die Verteilung jedoch kontinuierlich ist, darf das verteilungsfreie Verfahren verwendet werden. Am Ende des Dokumentes werden außerdem Hinweise auf ausgewählte wissenschaftliche Literatur gegeben, die die Ermittlung von Anteilsbereichen für andere Verteilungen unterstützen kann.

4 Procedures

4.1 Normal population with known variance and known mean

When the values of the mean, μ , and the variance, σ^2 , of a normally distributed population are known, the distribution of the characteristic under investigation is fully determined. There is exactly a proportion p of the population:

- a) to the right of $x_L = \mu - u_p \times \sigma$ (one-sided interval);
- b) to the left of $x_U = \mu + u_p \times \sigma$ (one-sided interval);
- c) between $x_L = \mu - u_{(1+p)/2} \times \sigma$ and $x_U = \mu + u_{(1+p)/2} \times \sigma$ (two-sided interval).

NOTE As such statements are known to be true, they are made with 100 % confidence.

In the above equations, u_p is p -fractile of the standard normal distribution. Numerical values of u_p may be read from the bottom line of the Tables B.1 to B.6 and Tables C.1 to C.6.

4.2 Normal population with known variance and unknown mean

Forms A and B, given in Annex A, are applicable to the case where the variance of the normal population is known while the mean is unknown. Form A applies to the one-sided case, while Form B applies to the two-sided case.

4.3 Normal population with unknown variance and unknown mean

Forms C and D, given in Annex A, are applicable to the case where both the mean and the variance of the normal population are unknown. Form C applies to the one-sided case, while Form D applies to the two-sided case.

4.4 Any continuous distribution of unknown type

If the characteristic under investigation is a continuous variable from a population of unknown form, and if a sample of n independent random observations of the characteristic has been taken, then a statistical tolerance interval can be determined from the ranked observations. The procedure given in Forms E and F of Annex A provide the determination of the coverage or sample size needed for tolerance intervals determined from the extreme values $x_{\min}$ or $x_{\max}$ of the sample of observations with given confidence level $1 - \alpha$.

NOTE Statistical tolerance intervals that do not depend on the shape of the sampled population are called *distribution-free* tolerance intervals.

This part of ISO 16269 does not provide procedures for distributions of known type other than the normal distribution. However, if the distribution is continuous, the distribution-free method may be used. Selected references to scientific literature that may assist in determining tolerance intervals for other distributions are also provided at the end of this document.

5 Beispiele

5.1 Daten

Formblätter A bis D in Anhang A werden durch Beispiele erläutert, die die Zahlenwerte aus ISO 2854:1976, Abschnitt 2, Absatz 1 der einleitenden Bemerkungen, Tabelle X, Garn 2, verwenden: 12 Messwerte der Bruchlast von Baumwollgarn. Es sollte beachtet werden, dass die Anzahl der hier für dieses Beispiel angegebenen Beobachtungen, $n = 12$, beträchtlich niedriger ist als die in ISO 2602 [1] empfohlene. Die Zahlenwerte und Berechnungen in den verschiedenen Beispielen sind in 10^{-2} Newton angegeben (siehe Tabelle 1).

Tabelle 1 — Daten für Beispiele 1 bis 4

Werte in 10^{-2} Newton

x	228,6	232,7	238,8	317,2	315,8	275,1	222,2	236,7	224,7	251,2	210,4	270,7
-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Diese Messergebnisse stammen aus einem Los von 12 000 Garnrollen aus einem Fertigungsauftrag, verpackt in 120 Kisten, jede mit 100 Garnrollen. Zwölf Kisten sind dem Los zufallsmäßig entnommen worden, und aus jeder dieser Kisten ist eine Garnrolle zufallsmäßig entnommen worden. In ungefähr 5 m Abstand vom freien Ende sind Teststücke von 50 cm Länge aus dem Garn auf diesen Garnrollen geschnitten worden. Die Messungen selbst sind an den Mittelstücken dieser Teststücke durchgeführt worden. Aufgrund vorliegender Informationen scheint es vernünftig anzunehmen, dass die unter diesen Bedingungen gemessenen Bruchlasten praktisch einer Normalverteilung folgen. In ISO 2954:1976 wird gezeigt, dass die Daten der Annahme einer Normalverteilung nicht widersprechen.

Aus diesen Messwerten folgt:

Stichprobenumfang: $n = 12$

Stichprobenmittelwert: $\bar{x} = 3\,024,1/12 = 252,01$

Stichprobenstandardabweichung: $s = \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}} = \sqrt{\frac{166\,772,27}{12 \times 11}} = \sqrt{1\,263,426\,3} = 35,545$

Die formale Darstellung der Berechnungen wird in Anhang A nur für Formblatt C (einseitig begrenzter Bereich, unbekannte Varianz) angegeben.

5.2 Beispiel 1: Einseitig begrenzter statistischer Anteilbereich bei bekannter Varianz

Aufgrund früherer Messergebnisse wird angenommen, dass die Streuung in Losen desselben Lieferanten konstant ist und durch eine Standardabweichung $\sigma = 33,150$ dargestellt wird, während der Erwartungswert nicht konstant ist. Es ist eine Grenze x_L zu ermitteln, für die bei einem Vertrauensniveau von $1 - \alpha = 0,95$ (95 %) gesagt werden kann, dass mindestens 0,95 (95 %) der unter den gleichen Bedingungen gemessenen Bruchlasten der Einheiten in dem Los über x_L liegen.

Tabelle B.4 liefert

$$k_1(12; 0,95; 0,95) = 2,120$$

woraus folgt

$$x_L = \bar{x} - k_1(n; p; 1 - \alpha) \times \sigma = 252,01 - 2,120 \times 33,150 = 181,732$$

5 Examples

5.1 Data

Forms A to D, given in Annex A, are illustrated by examples using the numerical values of ISO 2854:1976, Clause 2, paragraph 1 of the introductory remarks, Table X, yarn 2: 12 measures of the breaking load of cotton yarn. It should be noted that the number of observations, $n = 12$, given here for these examples is considerably lower than the one recommended in ISO 2602 [1]. The numerical data and calculations in the different examples are expressed in centi-newtons (see Table 1).

Table 1 — Data for Examples 1 to 4

Values in centi-newtons

x	228,6	232,7	238,8	317,2	315,8	275,1	222,2	236,7	224,7	251,2	210,4	270,7
-----	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

These measurements were obtained from a batch of 12,000 bobbins, from one production job, packed in 120 boxes each containing 100 bobbins. Twelve boxes have been drawn at random from the batch and a bobbin has been drawn at random from each of these boxes. Test pieces of 50 cm length have been cut from the yarn on these bobbins, at about 5 m distance from the free end. The tests themselves have been carried out on the central parts of these test pieces. Previous information makes it reasonable to assume that the breaking loads measured in these conditions have virtually a normal distribution. It is demonstrated in ISO 2954:1976 that the data do not contradict the assumption of a normal distribution.

These results yield the following:

Sample size: $n = 12$

Sample mean: $\bar{x} = 3\,024,1/12 = 252,01$

Sample standard deviation:
$$s = \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}} = \sqrt{\frac{166\,772,27}{12 \times 11}} = \sqrt{1\,263,426\,3} = 35,545$$

The formal presentation of the calculations will be given only for Form C in Annex A (one-sided interval, unknown variance).

5.2 Example 1: One-sided statistical tolerance interval under known variance

Suppose that previously obtained measurements have shown that the dispersion is constant from one batch to another from the same supplier, and is represented by a standard deviation $\sigma = 33,150$, although the mean is not constant. A limit x_L is required such that it is possible to assert with confidence level $1 - \alpha = 0,95$ (95 %) that at least 0,95 (95 %) of the breaking loads of the items in the batch, when measured under the same conditions, are above x_L .

Table B.4 gives

$$k_1(12; 0,95; 0,95) = 2,120$$

whence

$$x_L = \bar{x} - k_1(n; p; 1 - \alpha) \times \sigma = 252,01 - 2,120 \times 33,150 = 181,732$$

Für die untere Grenze x_L würde sich ein kleinerer Wert ergeben, wenn ein größerer Anteil der Grundgesamtheit (beispielsweise $p = 0,99$) und/oder ein höheres Vertrauensniveau (beispielsweise $1 - \alpha = 0,99$) gefordert würden.

5.3 Beispiel 2: Zweiseitig begrenzter statistischer Anteilsbereich bei bekannter Varianz

Unter denselben Bedingungen wie in Beispiel 1 werde angenommen, dass Grenzen x_L und x_U zu ermitteln sind, für die bei einem Vertrauensniveau von $1 - \alpha = 0,95$ gesagt werden kann, dass mindestens ein Anteil von $p = 0,90$ (90 %) der Bruchlasten des Loses zwischen x_L und x_U liegt.

Tabelle C.4 liefert

$$k_2(12; 0,90; 0,95) = 1,889$$

woraus folgt

$$x_L = \bar{x} - k_2(n; p; 1 - \alpha) \times \sigma = 252,01 - 1,889 \times 33,150 = 189,390$$

$$x_U = \bar{x} + k_2(n; p; 1 - \alpha) \times \sigma = 252,01 + 1,889 \times 33,150 = 314,630$$

Ein Vergleich mit Beispiel 1 zeigt, dass die Angabe, mindestens 90 % einer Grundgesamtheit liegen zwischen den Grenzen x_L und x_U , nicht das Gleiche ist wie die Angabe, nicht mehr als 5 % liegen außerhalb dieser Grenzen.

5.4 Beispiel 3: Einseitig begrenzter statistischer Anteilsbereich bei unbekannter Varianz

Hier werde angenommen, dass die Standardabweichung der Grundgesamtheit unbekannt ist und aus der Stichprobe geschätzt werden muss. Es seien die gleichen Forderungen wie in dem Fall bekannter Standardabweichung (Beispiel 1) angenommen, also $p = 0,95$ und $1 - \alpha = 0,95$. Die Ergebnisse sind im Einzelnen weiter unten dargestellt.

Ermittlung des statistischen Anteilsbereichs für den Anteil p :

a) einseitig nach unten begrenzter Anteilsbereich

Festgelegte Werte:

b) für den Anteilsbereich gewählter Anteil der Grundgesamtheit: $p = 0,95$

c) gewähltes Vertrauensniveau: $1 - \alpha = 0,95$

d) Stichprobenumfang: $n = 12$

Wert des Faktors aus Tabelle D.4:

$$k_3(n; p; 1 - \alpha) = 2,737$$