INTERNATIONAL STANDARD

ISO 16269-6

First edition
2005-04-01

Statistical interpretation of data —

Part 6: Determination of statistical tolerance intervals

Interprétation statistique des données —

Partie 6: Détermination des intervalles statistiques de tolérance

Reference number
ISO 16269-6:2005(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. 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.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 16269-6 was prepared by Technical Committee ISO/TC 69, *Applications of statistical methods*.

This first edition of ISO 16269-6 cancels and replaces ISO 3207:1975, which has been technically revised.

ISO 16269 consists of the following parts, under the general title *Statistical interpretation of data*:

— *Part 6: Determination of statistical tolerance intervals*

— *Part 7: Median — Estimation and confidence intervals*

— *Part 8: Determination of prediction intervals*

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].

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.

Statistical interpretation of data —

Part 6: **Determination of statistical tolerance intervals**

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.

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

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].

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 ith observed value $(i = 1, 2, \ldots, n)$

$x_{\max}$ maximum value of the observed values: $x_{\max} = \max\{x_1, x_2, \ldots, x_n\}$

$x_{\min}$ minimum value of the observed values: $x_{\min} = \min\{x_1, x_2, \ldots, x_n\}$

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

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

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

A smaller value of the lower limit x_L would be obtained if a larger proportion of the population (for example $p = 0{,}99$) and/or a higher confidence level (for example $1 - \alpha = 0{,}99$) were required.

5.3 Example 2: Two-sided statistical tolerance interval under known variance

Under the same conditions as in Example 1, suppose that limits x_L and x_U are required such that it is possible to assert with a confidence level $1 - \alpha = 0{,}95$ that at least a proportion of $p = 0{,}90$ (90 %) of the breaking load of the batch falls between x_L and x_U.

Table C.4 gives

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

whence

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

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

Comparison with Example 1 should make it clear that assuring that at least 90 % of a population lies between the limits x_L and x_U is not the same thing as assuring that no more than 5 % lies beyond each limit.

5.4 Example 3: One-sided statistical tolerance interval under unknown variance

Here, it is supposed that the standard deviation of the population is unknown and has to be estimated from the sample. The same requirements will be assumed as for the case where the standard deviation is known (Example 1), thus, $p = 0{,}95$ and $1 - \alpha = 0{,}95$. The presentation of the results is given in detail below.

Determination of the statistical tolerance interval of proportion p:

a) one-sided interval "to the right"

Determined values:

b) proportion of the population selected for the tolerance interval: $p = 0{,}95$

c) chosen confidence level: $1 - \alpha = 0{,}95$

d) sample size: $n = 12$

Value of tolerance factor from Table D.4:

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

Calculations:

$$\bar{x} = \sum x / n = 252{,}01$$

$$s = \sqrt{\frac{n\sum x^2 - \left(\sum x\right)^2}{n(n-1)}} = 35{,}545$$

$$k_3(n;\ p; 1-\alpha)\times s = 97{,}286\ 7$$

Results: one-sided interval "to the right"

The tolerance interval which will contain at least a proportion p of the population with confidence level $1 - \alpha$ has a lower limit

$$x_L = \bar{x} - k_3(n;\ p; 1-\alpha)\times s = 154{,}723$$

5.5 Example 4: Two-sided statistical tolerance interval under unknown variance

Under the same conditions as in Example 2, suppose it is required to calculate the limits x_L and x_U such that it is possible to assert with a confidence level $1 - \alpha = 0{,}95$ that in a proportion of the batch at least equal to $p = 0{,}90$ (90 %) the breaking load falls between x_L and x_U.

Table E.4 gives

$$k_4(n;\ p;\ 1-\alpha) = 2{,}671$$

whence

$$x_L = \bar{x} - k_4(n;\ p;\ 1-\alpha) \times s = 252{,}01 - 2{,}671 \times 35{,}545 = 157{,}069$$

$$x_U = \bar{x} + k_4(n;\ p;\ 1-\alpha) \times s = 252{,}01 + 2{,}671 \times 35{,}545 = 346{,}951$$

It will be noted that the value of x_L is smaller and the value of x_U higher than in Example 2 (known variance), because the use of s instead of σ requires a larger value of the tolerance factor to allow for the extra uncertainty. It is necessary to have to pay a penalty for not knowing the population standard deviation σ and the extension of the statistical tolerance interval takes this into account. Of course, it is not quite sure that the value $\sigma = 33{,}150$ used in Examples 1 and 2 is correct. Therefore, it is wiser to use the estimate, s, together with Tables D.4 or E.4.

5.6 Example 5: Distribution-free statistical tolerance interval for continuous distribution

In a fatigue test by rotational stress carried out on a component of an aeronautical engine, a sample of 15 items has given the results (measurement of endurance), shown in ascending order of values in Table 2.

Table 2 — Data for Example 5

x	0,200	0,330	0,450	0,490	0,780	0,920	0,950	0,970	1,040	1,710	2,220	2,275	3,650	7,000	8,800

A graphical examination of checking normality, such as probability plot, shows that the hypothesis of normality for the population of components should almost certainly be rejected (see ISO 5479). The methods of Form E, given in Annex A, for determination of a statistical tolerance interval are therefore applicable.

The extreme values from the sample of $n = 15$ measurements are:

$x_{\text{min}} = 0{,}200$, $x_{\text{max}} = 8{,}800$

Suppose that the required confidence level $1 - \alpha$ is 0,95.

a) What is the maximum proportion of the population of components that will fall below $x_{\text{min}} = 0{,}200$? Table F.1, for $1 - \alpha = 0{,}95$, gives for the minimum proportion above x_{min} a value of p slightly higher than 0,75 (75 %). Hence, for the maximum proportion below x_{min} a value of $1 - p$ slightly lower than 0,25 (25 %).

b) What sample size is necessary for it to be possible to assert, at a confidence level 0,95, that a proportion at least $p = 0{,}90$ (90 %) of the population of components will be found below the largest of the values from that sample? Table F.1, for $1 - \alpha = 0{,}95$ and $p = 0{,}90$, gives $n = 29$.