INTERNATIONAL STANDARD 3207

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Statistical interpretation of data — Determination of a statistical tolerance interval

Interprétation statistique des données — Détermination d'un intervalle statistique de dispersion

First edition — 1975-05-15

UDC 519 (083.4)

Ref. No. ISO 3207-1975 (E)

Descriptors : statistical analysis, statistical tolerance, tables (data), computation.

Price based on 15 pages

FOREWORD

ISO (the International Organization for Standardization) is a worldwide federation of national standards institutes (ISO Member Bodies). The work of developing International Standards is carried out through ISO Technical Committees. Every Member Body interested in a subject for which a Technical Committee has been set up 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.

Draft International Standards adopted by the Technical Committees are circulated to the Member Bodies for approval before their acceptance as International Standards by the ISO Council.

International Standard ISO 3207 was drawn up by Technical Committee ISO/TC 69, *Applications of statistical methods,* and circulated to the Member Bodies in November 1973.

It has been approved by the Member Bodies of the following countries :

Australia	India	South Africa, Rep. of
Belgium	Israel	Switzerland
Bulgaria	Italy	Thailand
Czechoslovakia	Netherlands	Turkey
France	New Zealand	United Kingdom
Germany	Poland	U.S.S.R.
Hungary	Romania	

The Member Bodies of the following countries expressed disapproval of the document on technical grounds :

Sweden
U.S.A.

Printed in Switzerland

CONTENTS

This page intentionally left blank

Statistical interpretation of data — Determination of a statistical tolerance interval

SECTION ONE : FORMAL PRESENTATION OF RESULTS

GENERAL REMARKS

1) This International Standard specifies methods enabling a sample to be used as the basis for determining a statistical tolerance interval, i.e. an interval such that there is a fixed probability (confidence level) that the interval will contain at least a proportion p of the population from which the sample is taken. The statistical tolerance interval may be two-sided or one-sided. The limits of the interval are called "statistical tolerance limits"; they are also called "natural limits of the process".

2) These methods are applicable only where it may be assumed that in the population under consideration the sample units have been selected at random and are independent.

3) The methods described below apply also only on condition that the distribution of the characteristic being studied is normal. The requirement of normality is more important here than for the inferences on means and differences between means in ISO 2854, *Statistical interpretation of data — Techniques of estimation and tests relating to means and variances.*

4) In order to check the hypothesis of normality, the methods laid down in ISO. . ., *Statistical interpretation of data — Normality tests*[1], are used.

5) Where the hypothesis of normality has to be rejected or where there is some reason to doubt its validity, one may envisage transforming the variate to make it normal or applying the method described in the introductory remark of annex A of this International Standard.

It is also possible to apply now methods which allow the determination of statistical tolerance intervals for other distribution forms than normal distributions. The description of these methods has not been considered in this International Standard.

6) In determining a statistical tolerance interval, it is desirable in connection with the origin or the method of collection of data to give all information that may assist in their statistical analysis, in particular the smallest unit or fraction of a measurement unit having practical significance.

7) No elimination or potential correction of individual data that are doubtful shall be carried out unless there are experimental, technical or obvious reasons to provide circumstantial justification of such elimination or correction.

In every instance, mention shall be made of the data eliminated or corrected.

8) As stated in 1), the confidence level $1-\alpha$ is the probability that the statistical tolerance interval will contain at least a proportion p of the population. The risk of this interval containing less than a proportion p of the population is α. The most usual values of $1-\alpha$ are 0,95 and 0,99 ($\alpha = 0{,}05$ and 0,01).

This means that if statistical tolerance intervals are determined for a large number of samples at the confidence level 0,95 for example, the proportion of those intervals which will contain at least the desired fraction of the population will be close to 95 %.

9) Tables 1 and 2 are applicable to the case where the standard deviation for the population is known (the mean being unknown); tables 3 and 4 to the case where the mean and the standard deviation are unknown.

Where the mean and the standard deviation having respectively the values m and σ are known, the distribution of the characteristic under investigation (assumed to be normal) is fully determined; there is exactly a proportion p of the population :

- on the right side of $m - u_p\sigma$ } one-sided intervals
- on the left side of $m + u_p\sigma$ }
- between $m - u_{(1+p)/2}\,\sigma$ and $m + u_{(1+p)/2}\,\sigma$: two-sided interval

where u_p is the fractile of order p of the standardized normal variate.

Numerical values of u_p are in these cases to be read on the bottom line of tables 5 and 6.

10) The calculations can often be very much simplified by making a change in origin and/or in unit.

1) In preparation.

TABLE 1 — **One-sided statistical tolerance interval** (known variance)[1]

Technical characteristics of the population under investigation[2] . . . .

Technical characteristics of the sample units[2] . . . .

Eliminated observations[3] . . . .

Statistical data	**Calculations**
Sample size : $n =$	
Sum of the observed values : $\Sigma x =$	$\bar{x} = \frac{\Sigma x}{n} =$
Known value of the variance of the population : $\sigma^2 =$	
whence the standard deviation : $\sigma =$	$k_1\ (n, p, 1-\alpha)\ \sigma =$ [6]
Proportion of the population selected for the statistical tolerance interval[4] : $p =$	
Chosen confidence level[5] : $1-\alpha =$	
$k_1\ (n, p, 1-\alpha) =$	

Results

a) One-sided interval "to the left"

There is a probability $1-\alpha$ that at least a proportion p of the population is above :

$$L_s = \bar{x} + k_1\ (n, p, 1-\alpha)\ \sigma =$$

b) One-sided interval "to the right"

There is a probability $1-\alpha$ that at least a proportion p of the population is above :

$$L_i = \bar{x} - k_1\ (n, p, 1-\alpha)\ \sigma =$$

1) A numerical example is given in section two of this International Standard : example No. 1
2) See paragraph 6 of General remarks.
3) See paragraph 7 of General remarks.
4) See paragraph 1 of General remarks.
5) See paragraph 8 of General remarks.
6) The values of $k_1\ (n, p, 1-\alpha)$ can be read directly from table 5 for different values of n, and for

$$p = 0{,}90;\ 0{,}95;\ 0{,}99$$

$$1-\alpha = 0{,}95 \text{ and } 0{,}99$$

TABLE 2 — **Two-sided statistical tolerance interval** (known variance)[1]

Technical characteristics of the population under investigation[2]	
Technical characteristics of the sample units[2]	
Eliminated observations[3]	
Statistical data	**Calculations**
Sample size : $n =$	
Sum of the observed values : $\Sigma x =$	$\bar{x} = \frac{\Sigma x}{n} =$
Known value of the variance of the population : $\sigma^2 =$	
whence the standard deviation : $\sigma =$	$k'_1\ (n, p, 1-\alpha)\ \sigma =$ [6]
Proportion of the population selected for the statistical tolerance interval[4] : $p =$	
Chosen confidence level[5] : $1-\alpha =$	
$k'_1\ (n, p, 1-\alpha) =$	
Results	
There is a probability $1-\alpha$ that at least a proportion p of the population is included between the limits[7] :	
$L_i = \bar{x} - k'_1\ (n, p, 1-\alpha)\ \sigma =$	
$L_s = \bar{x} + k'_1\ (n, p, 1-\alpha)\ \sigma =$	

1) A numerical example is given in section two of this International Standard : example No. 2.
2) See paragraph 6 of General remarks.
3) See paragraph 7 of General remarks.
4) See paragraph 1 of General remarks.
5) See paragraph 8 of General remarks.
6) The values of $k'_1\ (n, p, 1-\alpha)$ can be read from table 6 for different values of n, and for

$$p = 0{,}90;\ 0{,}95;\ 0{,}99$$

$$1-\alpha = 0{,}95 \text{ and } 0{,}99$$

7) These limits are symmetrical about $\bar{x}$ but they are not "symmetrical in probability". It is not true that at the confidence level $1-\alpha$, a proportion not exceeding $(1-p)/2$ of the population is below L_i and a proportion not exceeding $(1-p)/2$ is above L_s.

TABLE 3 — **One-sided statistical tolerance interval** (unknown variance)[1]

Technical characteristics of the population under investigation[2] . . . .

Technical characteristics of the sample units[2] . . . .

Eliminated observations[3] . . . .

Statistical data	**Calculations**
Sample size : $n =$	$\overline{x} = \dfrac{\Sigma x}{n} =$
Sum of the observed values : $\Sigma x =$	$\dfrac{\Sigma (x - \overline{x})^2}{n-1} = \dfrac{\Sigma x^2 - (\Sigma x)^2/n}{n-1} =$
Sum of the squares of the observed values : $\Sigma x^2 =$	
Proportion of the population selected for the statistical tolerance interval[4] : $p =$	$\sigma^* = s = \sqrt{\dfrac{\Sigma (x - \overline{x})^2}{n-1}} =$ (estimation of the standard deviation σ)
Chosen confidence level[5] : $1 - \alpha =$	$k_2 \, (n, p, 1 - \alpha) \, s =$
$k_2 \, (n, p, 1 - \alpha) =$ [6]	

Results

a) One-sided interval "to the left"

There is a probability $1 - \alpha$ that at least a proportion p of the population is below :

$$L_s = \overline{x} + k_2 \, (n, p, 1 - \alpha) \, s =$$

b) One-sided interval "to the right"

There is a probability $1 - \alpha$ that at least a proportion p of the population is above :

$$L_i = \overline{x} - k_2 \, (n, p, 1 - \alpha) \, s =$$

1) A numerical example is given in section two of this International Standard : example No. 3.
2) See paragraph 6 of General remarks.
3) See paragraph 7 of General remarks.
4) See paragraph 1 of General remarks.
5) See paragraph 8 of General remarks.
6) The values of $k_2 \, (n, p, 1 - \alpha)$ can be read from table 7 for different values of n, and for

$$p = 0{,}90;\ 0{,}95;\ 0{,}99$$

$$1 - \alpha = 0{,}95 \text{ and } 0{,}99$$

TABLE 4 — **Two-sided statistical tolerance interval** (unknown variance)[1]

Technical characteristics of the population under investigation[2] . . . .

Technical characteristics of the sample units[2] . . . .

Eliminated observations[3] . . . .

Statistical data	**Calculations**
Sample size : $n =$	$\overline{x} = \frac{\Sigma x}{n} =$
Sum of the observed values : $\Sigma x =$	$\frac{\Sigma (x - \overline{x})^2}{n-1} = \frac{\Sigma x^2 - (\Sigma x)^2/n}{n-1} =$
Sum of the squares of the observed values : $\Sigma x^2 =$	
Proportion of the population selected for the statistical tolerance interval[4] : $p =$	$\sigma^* = s = \sqrt{\frac{\Sigma (x - \overline{x})^2}{n-1}} =$ (estimation of the standard deviation σ)
Chosen confidence level[5] : $1 - \alpha =$	
$k'_2\ (n, p,\ 1 - \alpha) =$	$k'_2\ (n, p,\ 1 - \alpha)\, s =$ [6]

Results

There is a probability $1 - \alpha$ that at least a proportion p of the population is included between the limits[7] :

$$L_i = \overline{x} - k'_2\ (n, p,\ 1 - \alpha)\, s =$$

$$L_s = \overline{x} + k'_2\ (n, p,\ 1 - \alpha)\, s =$$

1) A numerical example is given in section two of this International Standard : example No. 4.
2) See paragraph 6 of General remarks.
3) See paragraph 7 of General remarks.
4) See paragraph 1 of General remarks.
5) See paragraph 8 of General remarks.
6) The values of $k'_2\ (n, p,\ 1 - \alpha)$ can be read from table 8 for different values of n, and for

$$p = 0{,}90;\ 0{,}95;\ 0{,}99$$
$$1 - \alpha = 0{,}95 \text{ and } 0{,}99$$

7) These limits are symmetrical about $\bar{x}$ but they are not "symmetrical in probability". It is not true that at the confidence level $1 - \alpha$, a proportion not exceeding $(1 - p)/2$ of the population is below L_i and a proportion not exceeding $(1 - p)/2$ is above L_s.

SECTION TWO : EXAMPLES

INTRODUCTORY REMARKS

Tables 1 to 4 will be illustrated by examples using the numerical values of ISO 2854 (section two, paragraph 1 of the introductory remarks, table X, yarn 2) : 12 measurements 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 2062, *Textiles — Yarn from packages — Method for determination of breaking load and elongation at the breaking load of single strands — (CRL, CRE and CRT testers).*

The unit of measurement used to express the numerical data and the results of calculations in the different examples is the centinewton.

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 come from a batch of 12 000 bobbins, from one production job, packed in 120 boxes each containing 100 bobbins. 12 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 enables it to be allowed that the breaking loads measured in these conditions have virtually a normal distribution..

These results yield the following :

Sample size :

$$n = 12$$

Sum of the observed values :

$$\Sigma x = 3\,024{,}1$$

Mean :

$$\overline{x} = \frac{\Sigma x}{n} = 252{,}0$$

Sum of the squares of the observed values :

$$\Sigma x^2 = 775\,996{,}09$$

Sum of the squares of the differences from the mean value :

$$\Sigma\,(x - \overline{x})^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 13\,897{,}69$$

Estimated variance :

$$s^2 = \frac{\Sigma\,(x - \overline{x})^2}{n - 1} = 1\,263{,}4$$

Estimated standard deviation :

$$s = 35{,}5$$

It is also known from experience that in the same batch the breaking loads are distributed according to a pattern very close to normal distribution.

The formal presentation of the calculations will be given only for table 3 (one-sided interval, unknown variance).

NUMERICAL EXAMPLES

Example No. 1 — **One-sided statistical tolerance interval** (known variance, table 1)

One assumes that the measurements previously obtained have shown that the dispersion is constant from one batch to another from the same supplier, although the mean is not constant, and is represented by a standard deviation $\sigma = 33{,}15$.

One wishes to calculate the limit L_i such that it is possible to assert with a confidence level $1 - \alpha = 0{,}95$ that in a proportion at least equal to 0,95 (95 %) the breaking load of the items likely to be taken as samples from the batch and measured in the same conditions is above L_i.

Table 5 gives :

$$k_1\ (12;\ 0{,}95;\ 0{,}95) = 2{,}12$$

whence :

$$L_i = \overline{x} - k_1 \sigma = 252{,}0 - 2{,}12 \times 33{,}15 = 181{,}7$$

One would, of course, obtain a smaller limit L_i if one took a higher proportion of the population (for example 99 %) and/or a higher confidence level (for example $1 - \alpha = 0{,}99$).

Example No. 2 — **Two-sided statistical tolerance interval** (known variance, table 2)

In the same conditions as in example No. 1, one wishes to calculate the limits L_i and L_s 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 L_i and L_s.

Table 6 gives :

$$k'_1\ (12;\ 0{,}90;\ 0{,}95) = 1{,}89$$

whence :

$$L_i = \overline{x} - k'_1\ \sigma = 252{,}0 - 1{,}89 \times 33{,}15 = 189{,}3$$

$$L_s = \overline{x} + k'_1\ \sigma = 252{,}0 + 1{,}89 \times 33{,}15 = 314{,}7$$

An attempt must be made to eliminate the suspicion that 5 % (at the maximum) of the population lies to the left of L_i and similarly, to the right of L_s.

It has been seen, in connection with example No. 1, that the limit to the left of which not more than a maximum of 5 % of the population lies (minimum of 95 % to the right of L_i) is $L_i = 181,7$.

Example No. 3 — **One-sided statistical tolerance interval** (unknown variance, table 3)

Here, one assumes that the standard deviation of the population is unknown and has to be estimated from the sample. One will take the same conditions as for the case where the standard deviation is known (example No. 1) — thus, $p = 0,95$ and $1 - \alpha = 0,95$.

The presentation of the results is given in detail below.

Technical characteristics of the population under investigation — The batch comprises a delivery of cotton yarn received on 1969-08-03 from the supplier H and consisting of 12 000 bobbins packed in 120 boxes each containing 100 bobbins.

Technical characteristics of the sample units — 12 boxes have been drawn at random from the batch and one bobbin has been taken 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.

Eliminated observations : none

Statistical data

Sample size :

$n = 12$

Sum of the observed values :

$\Sigma x = 3\,024,1$

Sum of the squares of the observed values :

$\Sigma x^2 = 775\,996,09$

Proportion of the population selected for the statistical tolerance interval :

$p = 0,95$ (95 %)

Chosen confidence level :
$1 - \alpha = 0,95$
k_2 (12; 0,95; 0,95) = 2,74
(value read from table 7)

Calculations

$$\overline{x} = \frac{3\,024,1}{12} = 252,0$$

$$\frac{\Sigma (x - \overline{x})^2}{n-1} = \frac{\Sigma x^2 - (\Sigma x)^2/n}{n-1}$$

$$= \frac{775\,996,09 - (3\,024,1)^2/12}{11}$$

$$= 1\,263,4$$

$$\sigma^* = s = \sqrt{1\,263,4} = 35,5$$

$$k_2\ (12; 0,95; 0,95)\ s = 97,3$$

Results

It is possible to assert with a confidence level of 0,95 that a proportion at least equal to 0,95 (95 %) of the breaking loads for the batch is above :

$$L_i = 252,0 - 2,74 \times 35,5 = 154,7$$

It will be noted that the value of L_i is smaller than in example No. 1 (known variance), because the use of s, random estimation of σ, produces a higher value of the coefficient k (2,74 instead of 2,12).