*Indian Standardभारतीय मानक*

> **परीक्षण के पररणामों का साांख्यिकीि ख्िर्वचि — माध्ि, मािक ख्र्चलि और समाश्रिण ग ु णाांक का आांकलि — ख्र्श्वास्िता अांतराल** *( पहला पनरीक्षण ु )*

**Statistical Interpretation of Test Results — Estimation of Mean, Standard Deviation and Regression Coefficient — Confidence Interval** 

*( First Revision )* 

ICS 03.120.30

© BIS 2024



भारतीय मानक ब्यरोू BUREAU OF INDIAN STANDARDS मानक भवन, 9 बहादुर शाह ज़फर मार्ग, नई दिल्ली -  $110002$ MANAK BHAVAN, 9 BAHADUR SHAH ZAFAR MARG NEW DELHI - 110002 [www.bis.gov.in](http://www.bis.org.in/) [www.standardsbis.in](http://www.standardsbis.in/)

**June 2024 Price Group 7** 

<span id="page-1-0"></span>Statistical Methods for Quality, Data Analytics and Reliability Sectional Committee, MSD 03

## FOREWORD

This Indian Standard (First Revision) was adopted by the Bureau of Indian Standards, after the draft finalized by the Statistical Methods for Quality, Data Analytics and Reliability Sectional Committee had been approved by the Management and Systems Division Council.

Industrial experimentation is often performed to estimate some unknown values. These values are usually parameters (constants, such as, mean thickness of metal sheets; variance of tensile strength of wire; relationship between carbon content and tensile strength of steel, etc) of a probability distribution or function of these parameters. The present standard deals with estimation of parameters of a normal population on the basis of series of test results on items in a sample drawn from this population, and discusses point estimates and interval estimates for mean, standard deviation and regression coefficient.

A point estimate is a single value which is used to estimate the parameter in question. A point estimate is often inadequate as an estimate of a parameter since it rarely coincides with the value of the parameter and does not indicate how far away this estimate is from the true value of the parameter. One way of expressing this uncertainty is to specify an interval as an estimate instead of a single value, stating that the interval thus calculated includes the true value of the population parameter, has a specified high probability. Such an interval is called a confidence interval.

The confidence interval is obtained as a function of the test results on items in the sample of observations and therefore is a random interval. Associated with it is a confidence level (sometimes termed as confidence coefficient), which is the probability, usually expressed as a percentage, that the interval does contain the parameter of the population. If we choose an interval such that the probability that it contains the value of the population parameter is  $1 - \alpha$ , then we say that the interval is a 100 ( $1 - \alpha$ ) percent confidence interval for the parameter and 1 - *α* is known as the confidence level. Generally, 95 percent and 99 percent confidence intervals are constructed.

This standard was first published in 1996. The first revision has been made in the light of experience gained since its publication and to incorporate the following major changes:

- a) Two new tables have been added to give relevant values of standardized *t and*  $\chi^2$  distribution;
- b) Align it with IS/ISO 16269-7 'Statistical interpretation of data: Part 7 Median Estimation and confidence intervals';
- c) Point estimate of mean and standards deviation for grouped data as in ISO 2602 : 1980 'Statistical interpretation of test results — Estimation of the mean — Confidence interval' has been added; and
- d) Estimation of confidence interval of mean using range has also been added.

The composition of the Committee responsible for the formulation of this standard is given in [Annex B](#page-11-0).

For the purpose of deciding whether a particular requirement of this standard is complied with, the final value, observed or calculated, expressing the result of a test or analysis, shall be rounded off in accordance with IS 2 : 2022 'Rules for rounding off numerical values (*second revision*)'. The number of significant places retained in the rounded off value should be the same as that of the specified value in this standard.

## *Indian Standard*

# <span id="page-2-0"></span>STATISTICAL INTERPRETATION OF TEST RESULTS — ESTIMATION OF MEAN, STANDARD DEVIATION AND REGRESSION COEFFICIENT — CONFIDENCE INTERVAL

## *( First Revision )*

## **1 SCOPE**

**1.1** This standard specifies the statistical treatment of test results needed to calculate confidence interval for the mean, standard deviation and regression coefficient of a population.

**1.2** It does not cover the calculation of an interval with a fixed probability, at least a given percentage of the population (that is, statistical tolerance interval) which is covered in IS/ISO 16269-6.

## **2 REFERENCES**

The standard given below contain provisions which through reference in this text, constitute provision of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of these standard:



## **3 DEFINITIONS**

For the purpose of this standard, the definitions given in IS 7920 (Part 1) shall apply.

## **4 ESTIMATION OF MEAN**

## **4.1 Ungrouped Data**

A point estimate of the population mean (designated as  $\mu$ ) is given by the sample mean  $(\bar{x})$  and is obtained as follows:

$$
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i
$$

where  $x_1, x_2, \ldots, x_n$  are test results of *n* items in the sample.

#### **4.2 Grouped Data**

When the number of results is sufficiently high (for example, above 50), it may be advantageous to group them into classes of the same width. In certain cases, the results may also have been directly obtained grouped into classes.

The frequency of the  $i<sup>th</sup>$  class, that is, the number of results in class *i*, is denoted by *fi*.

The number of classes being denoted by *k*, we have:

$$
n = \sum_{i=1}^{k} f_i
$$

The midpoint of class  $i$  is designated by  $x_i$ . The mean  $\bar{x}$  is then estimated by the weighted mean of all midpoints of classes:

$$
\bar{x} = \frac{1}{n} \sum_{i=1}^{k} f_i x_i
$$

#### **5 ESTIMATION OF THE STANDARD DEVIATION**

## **5.1Ungrouped Results**

The point estimate of the population standard deviation  $(\sigma)$ , is given by sample standard deviation (*s*) as follows:

$$
s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}
$$

or

$$
s = \sqrt{\frac{1}{n-1} \left[ \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n} \right]}
$$

To access Indian Standards click on the link below:

[https://www.services.bis.gov.in/php/BIS\\_2.0/bisconnect/knowyourstandards/Indian\\_standards/isdetails/](https://www.services.bis.gov.in/php/BIS_2.0/bisconnect/knowyourstandards/Indian_standards/isdetails/)

<span id="page-3-0"></span>where

$$
x_i
$$
 = the value of the *i*<sup>th</sup> test result (*i* =1, 2, ...,   
n);

- $n =$  the total number of test results; and
- $\bar{x}$  = the sample mean calculated as in  $\frac{4.1}{x}$  $\frac{4.1}{x}$  $\frac{4.1}{x}$ .

## **5.2 Grouped Results in Classes**

In the case of grouping by classes, the formula for the estimate of the standard deviation is written:

$$
s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{k} f_i (x_i - \bar{x})^2}
$$

or

$$
= \sqrt{\frac{1}{n-1} \left[ \sum_{i=1}^{k} f_i x_i^2 - \frac{(\sum_{i=1}^{n} f_i x_i)^2}{n} \right]}
$$

 $\mathsf{l}$ 

where

 $\overline{S}$ 

 $f_i$  = frequency of  $i^{\text{th}}$  class;

 $x_i$  = the midpoint of  $i$ <sup>th</sup> class ( $i$  =1, 2, ..., k);

 $\bar{x}$  = the sample mean calculated as in [4.2](#page-2-0); and

 $n = \sum_{i=1}^{k} f_i$  = the total number of test results.

## **6 ESTIMATION OF THE REGRESSION COEFFICIENT**

Consider the linear regression of an independent variable (*X*) on the dependent variable (*Y*),

$$
E(Y/X = x) = \beta_0 + \beta_{1x}
$$

Given  $n$  (> 2) pairs of test results  $(x_i, y_i)$ ,  $i = 1, 2,...,$ *n*, the estimate of  $\beta_1$ , is:

$$
b_1 = \frac{s_{xy}}{s_x^2}
$$

where

 $s_{xy}$  = the estimate of covariance of *X* and *Y;* and

$$
s_x
$$
 = the estimate of standard deviation of X.

 $s_x$  is given by the expression in  $\overline{5}$  $\overline{5}$  $\overline{5}$  and  $s_{xy}$  is given by

the following expression:

$$
s_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})
$$

$$
s_{xy} = \frac{1}{n-1} \left[ \sum x_i y_i - \frac{\left( \sum x_i \sum y_i \right)}{n} \right]
$$

## **7 CONFIDENCE INTERVAL FOR THE MEAN**

#### **7.1Standard Deviation Known**

#### **7.1.1** *Two Sided*

The confidence interval for the population mean is calculated from the point estimate  $(\bar{x})$  of the mean and the known standard deviation (σ)*.*

A 100 (1 **-** *α*) percent confidence interval for the population mean of a normally distributed variable with mean  $\mu$  and variance  $\sigma$  *is* given by:

$$
\left[\; \bar{x} - z_{\left(1\right. -\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}, \qquad \bar{x} + \; z_{\left(1\right. -\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \right]
$$

where

 $z_1 - \alpha/2$  = the value of standard normal deviate with the area of  $(1 - \frac{\alpha}{2})$  to its left; and  $\frac{1}{2}$ 

$$
\frac{\sigma}{\sqrt{n}} = \text{ the standard error of } \bar{x}.
$$

**NOTES** 

**1** The value of  $z_1$ -  $\alpha$  /2 can be obtained from the tables of the standard normal. For 95 percent and 99 percent confidence intervals, the value is 1.96 and 2.575, respectively.

 $2$  As  $1 - \alpha$  increases, the length of the interval also increases. Also, for a fixed  $\alpha$ , as  $\sigma$  increases the length of the interval increases.

**3** For given α and *σ,* larger the sample size, narrower is the length of the interval and better the estimate.

#### **7.1.2** *One Sided*

In some cases, one-sided confidence intervals are most appropriate to use. The one-sided confidence interval may be either an upper limit or a lower limit. A 100  $(1 - \alpha)$  percent one-sided confidence interval for population mean of a normally distributed variable with mean (u) and variance (*a*) is given by:

$$
\left[\bar{x} - z_{(1-\alpha)}\frac{\sigma}{\sqrt{n}}, +\infty\right]
$$

$$
\quad \text{or} \quad
$$

$$
\[-\infty, \ \bar{x} + z_{(1-\alpha)}\frac{\sigma}{\sqrt{n}}\]
$$

<span id="page-4-0"></span>where

 $z_1$  -  $\alpha$  = the value of standard normal deviate with the area of  $(1 - \alpha)$  to its left and

σ  $\sqrt{n}$ = the standard error of  $\bar{x}$ .

## **7.2 Standard Deviation Not Known**

## **7.2.1** *Sample Size More Than* 30

When the parameter  $\sigma$  is not known and if the sample is large, say, greater than 30, its point estimate (*s*), as given in  $\overline{5}$  $\overline{5}$  $\overline{5}$ , is made use of in place of  $\sigma$  for calculating confidence interval as given in **[7.1](#page-3-0)**.

#### **7.2.2** *Sample Size Less Than Equal to* 30

If the sample size is less than equal to 30, its point estimate (*s*), as given in **5** is made use of in place of *σ*, and student's *t*-distribution is made use of in place of normal distribution for calculating confidence interval.

The 100  $(1 - \alpha)$  percent confidence interval is given by:

$$
\bar{x}-t_{1-\tfrac{\alpha}{2}(n-1)}\tfrac{s}{\sqrt{n}}\,,x+t_{1-\tfrac{\alpha}{2}(n-1)}\tfrac{s}{\sqrt{n}}
$$

For one-sided confidence interval a 100 (1- $\alpha$ ) percent confidence interval is given by:

$$
\bar{x} - t_{1-\alpha(n-1)} \frac{s}{\sqrt{n}}, +\infty
$$
  

$$
-\infty, \bar{x} - t_{1-\alpha(n-1)} \frac{s}{\sqrt{n}}
$$

The values of  $t_{1-\alpha}$  and the ratio  $t_{1-\alpha}/\sqrt{n}$  are given in the  $Table 1$  and  $Table 2$  respectively, for 95 percent and 99 percent confidence level for *n* - 1 degrees of freedom.

## *Example* 1

Measurements on the thickness (in codified

units) of a sample of 16 mica discs from a production process are as follows:

14, 11, 11, 17, 15, 13, 14, 11, 14, 12, 10, 10, 8, 13, 7, 8

The mean of thickness values is calculated as 11.75. Suppose it is known that these discs have been produced by a controlled process whose standard deviation is known to be equal to 2.5.

Following the procedure described in **[7.1](#page-3-0)**, we find that the 95 percent confidence interval for the mean is:

$$
\left[11.75 - \frac{1.96 \times 2.5}{4}, 11.75 + \frac{1.96 \times 2.5}{4}\right]
$$

Which gives us the interval [10.5, 13.0]. The probability that the population mean is included in the estimated interval [10.5, 13.0] is 95 percent. Similarly, the 99 percent confidence interval can be calculated as [10.1, 13.4].

#### *Example* 2

Consider the above example when there is no knowledge about the standard deviation of the thickness. The estimate of  $\sigma$  is obtained as  $s = 2.77$ .

In this case the 95 percent confidence interval is seen to be (values of *t* and  $\frac{t}{\sqrt{n}}$  for  $n = 16$  are obtained from [Table 1](#page-5-0) and [Table 2\)](#page-6-0).

$$
[11.75 - (2.131)\frac{s}{\sqrt{n}}, 11.75 + (2.131)\frac{s}{\sqrt{n}}]
$$

 $=$  [11.75 - 0.533  $\times$  2.77, 11.75 + 0.533  $\times$  2.77]

Which is [10.3, 13.2].

Similarly, the 99 percent confidence interval can be, calculated as [9.7, 13.8].

<span id="page-5-0"></span>

## **Table 1 Values of** *t***1 - α**

(*Clause* [7.2.2\)](#page-4-0)

<span id="page-6-0"></span>

## **Table 2 The Ratio**  $t_1 \cdot \alpha / \sqrt{n}$

(*Clause* [7.2.2\)](#page-4-0)

5

## <span id="page-7-0"></span>**7.3 Range**

Assuming that the population is normally distributed, the confidence interval for the population mean can be determined from the sample range when the number of measurements is small, say 10 or less. The practical convenience of this calculation is that it is faster; its disadvantage is that it leads to a confidence interval which is generally, wider and which is more sensitive to departures from the assumed normal form of the observations. Detailed formulae and calculations with example is given in [Annex A.](#page-10-0) 

## **8 CONFIDENCE INTERVAL FOR VARIANCE**

The one-sided confidence interval for variance  $\sigma^2$  is defined by the upper limit, the lower limit being taken as equal to zero.

The interval is given by:

$$
\left[0,\frac{(n-1)s^2}{\chi^2_{1-\alpha(n-1)}}\right]
$$

The values of  $\chi^2_{1-\alpha(n-1)}$  are given in the Table 3 for  $\alpha$  equal to 0.05 and 0.01.



## Table 3 Critical Values of  $\chi^2$ -Distribution

(*Clause* 8)

**Table 3** (*Concluded*)



## *Example* 3

The precision of a micrometer is measured in terms of the standard deviation of the readings made by it. The 10 readings made on a specimen with a calibrated micrometer are as follows:

0.501, 0.502, 0.498, 0.499, 0.501, 0.503, 0.499, 0.502, 0.497, 0.504

We are required to find the 95 percent confidence interval for the variance,  $\sigma^2$ , of these readings.

The calculated value of  $(n - 1) s^2$  is 0.000 046 4 while the value of  $\chi$  for  $n - 1 = 9$  is 16.92 for confidence level of 95 percent.

The one-sided 95 percent confidence interval is [0, 0.000 002 742].

## **9 CONFIDENCE INTERVAL FOR THE REGRESSION COEFFICIENT**

The confidence interval for  $b_1$  is given by:

$$
b_1 - t_{1-\alpha/2(n-2)} \sqrt{\frac{s_y^2 - b_1 s_{xy}}{(n-2)s_x^2}}
$$
  

$$
b_1 + t_{1-\alpha/2(n-2)} \sqrt{\frac{s_y^2 - b_1 s_{xy}}{(n-2)s_x^2}}
$$

where

$$
b_1, s_x = \text{as defined in 6; and}
$$
  
and 
$$
s_{xy}
$$

 $s_y$  = the sample standard deviation of *y* as defined for  $\overline{x}$  in  $\underline{6}$  $\underline{6}$  $\underline{6}$ .

 $(n-2)s_x^2$ 

## **IS 14277 : 2024**

#### *Example* 4

The temperature in the heating zone of an exhaust is related to the duration of heating. To estimate this relationship readings were taken on the temperature once every 3 minutes beginning 2 minutes after the heating started. These readings were as follows:



We take

 $x =$  duration in minutes; and

 $y =$  temperature in  $\mathrm{C}$ .

The point-estimate of the regression coefficient  $\beta_1$ , that is,  $b_1$  is equal to 6.12.

We also have  $s_{xy} = 4545$  and  $s_x^2 = 742.5$  and  $s_y^2$  $= 28 090.$ 

Thus, the 95 percent confidence interval is obtained as:

$$
\[6.12 - (2.306) \times \sqrt{\frac{33.64}{742.5}}\]
$$

$$
6.12 + (2.306) \times \sqrt{\frac{33.64}{742.5}}
$$

 $= [ 6.12 - 2.306 \times 0.213, 6.12 + 2.306 \times 0.213 ]$ 

This gives us the interval [5.6, 6.6].

The regression coefficient in this example measures the rate of change in temperature per minute increase in duration of heating. There are 95 percent chances that this rate is between 5.6 °C and 6.6 °C.

## **10 ROUNDING OFF**

The final values obtained after calculations shall be rounded off as per [IS 2](#page-2-0).

## **ANNEX A**

## (*Clause* [7.3](#page-7-0))

## **CONFIDENCE INTERVAL FOR THE MEAN FROM THE RANGE**

## <span id="page-10-0"></span>**A-1 TWO-SIDED CONFIDENCE INTERVAL**

The two-sided confidence interval for the population mean is defined by the following double inequality:

a) At the confidence level 95 percent:

 $\bar{x}$  –  $q_{0.975}$ <sup>r</sup> < *m* < *x* +  $q_{0.975}$ <sup>r</sup>

b) At the confidence level 99 percent:

$$
\bar{x} - q_{0.995}^{\rm r} < m < x + q_{0.995}^{\rm r}
$$

## **A-2 ONE-SIDED CONFIDENCE INTERVAL**

The one-sided confidence interval for the population mean is defined by one or other of the following inequalities:

a) At the confidence level 95 percent:

$$
m < \bar{x} + \mathbf{q}_{0.95}^{\mathrm{r}}
$$

or

 $m > \bar{x} - q_{0.95}^{11}$ 

b) At the confidence level 99 percent:

 $m < \bar{x} + \mathbf{q}_{0.99}$ <sup>r</sup> or

 $m > \bar{x} - q_{0.99}$ <sup>r</sup>

The coefficients  $q_{0.975}$ ,  $q_{0.995}$ ,  $q_{0.95}$   $q_{0.99}$  are given in Table 4.

## **Table 4 Coefficients Q0.975, Q0.995, Q0.95 Q0.99 Values**

(*Clause* A-2)



## **ANNEX B**

## (*[Foreword](#page-1-0)*)

## **COMMITTEE COMPOSITION**

<span id="page-11-0"></span>Statistical Methods for Quality, Data Analytics and Reliability Sectional Committee, MSD 03

Indian Statistical Institute, Kolkata PROF BIMAL K. ROY **(***Chairperson***)** Automotive Research Association of India (ARAI), Pune Bharat Electronics Ltd, Ghaziabad MS EKTA BHARDWAJ

Bureau of Indian Standards, New Delhi SHRIMATI SNEH LATA BHATNAGAR

CSIR - National Physical Laboratory, New Delhi DR JIJI T. J. PULIKKOTIL

Indian Statistical Institute (ISI), Chennai MS SUSHMA BENDRE

M.S. University of Baroda, Deptt. of Statistics, Vadodra

NABL/QCI, New Delhi SHRI N. VENKATESWARAN

National Institution of Medical Statistics (NIMS), ICMR, New Delhi

- In Personal Capacity (*B-279, Derawal Nagar, Delhi - 110009*)
- In Personal Capacity (*804, Chelsea Tower, Omaxe Heights Vibhuti Khand, Gomti Nagar, Lucknow - 226010*)
- In Personal Capacity (*D-49 Arya Nagar Apartment I.P. Extension New Delhi - 110092*)
- In Personal Capacity [*C-45, Kendriya Vihar, Sector 56, (Near Huda Market), Gurugram - 122011*]

#### *Organization Representative(s)*

SHRI P. P. DAMBAL

MS ANJALI SHARMA (*Alternate*)

FICCI, New Delhi SHRI S. C. ARORA SHRI MRITYUNJAY KUMAR (*Alternate*)

PROF K. MURALIDHARAN

SHRI AVIJIT DAS (*Alternate* I) MS ANITA RANI (*Alternate* II)

SHRI H. K. CHATURVEDI

RITES, Gurugram SHRI MANISH BHATNAGAR SHRI RAKESH KUMAR (*Alternate*)

DR. V. K. BHATIA

PROF S. CHAKRABORTY

SHRI P. K. GAMBHIR

SHRI L. K. MEHTA

BIS Directorate General SHRI ANUJ SWARUP BHATNAGAR, SCIENTIST 'G'/SENIOR DIRECTOR AND HEAD (MANAGEMENT AND SYSTEMS) [ REPRESENTING DIRECTOR GENERAL (*Ex-officio*)]

> *Member Secretary* SHRI ASHISH V. UREWAR SCIENTIST 'C'/DEPUTY DIRECTOR (MANAGEMENT AND SYSTEMS), BIS

Tris Page rass degradationally left bank

## **Bureau of Indian Standards**

BIS is a statutory institution established under the *Bureau of Indian Standards Act*, 2016 to promote harmonious development of the activities of standardization, marking and quality certification of goods and attending to connected matters in the country.

## **Copyright**

**Headquarters:**

BIS has the copyright of all its publications. No part of these publications may be reproduced in any form without the prior permission in writing of BIS. This does not preclude the free use, in the course of implementing the standard, of necessary details, such as symbols and sizes, type or grade designations. Enquiries relating to copyright be addressed to the Head (Publication & Sales), BIS.

## **Review of Indian Standards**

Amendments are issued to standards as the need arises on the basis of comments. Standards are also reviewed periodically; a standard along with amendments is reaffirmed when such review indicates that no changes are needed; if the review indicates that changes are needed, it is taken up for revision. Users of Indian Standards should ascertain that they are in possession of the latest amendments or edition by referring to the websitewww.bis.gov.in or www.standardsbis.in.

This Indian Standard has been developed from Doc No.: MSD 03 (20963).

## **Amendments Issued Since Publication**



## **BUREAU OF INDIAN STANDARDS**



**Branches** : AHMEDABAD, BENGALURU, BHOPAL, BHUBANESHWAR, CHANDIGARH, CHENNAI, COIMBATORE, DEHRADUN, DELHI, FARIDABAD, GHAZIABAD, GUWAHATI, HARYNA, HUBLI, HYDERABAD, JAIPUR, JAMMU & KASHMIR, JAMSHEDPUR, KOCHI, KOLKATA, LUCKNOW, MADURAI, MUMBAI, NAGPUR, NOIDA, PARWANOO, PATNA, PUNE, RAIPUR, RAJKOT, SURAT, VIJAYAWADA.