

Guide to

Statistical interpretation of data —

Part 1: Routine analysis of quantitative
data

Confirmed
October 2008

Committees responsible for this British Standard

The preparation of this British Standard was entrusted by the Quality, Management and Statistics Standards Policy Committee (QMS/-) to Technical Committee QMS/12, upon which the following bodies were represented:

- Association for Consumer Research (ACRE)
- Clay Pipe Development Association Limited
- Institute of Quality Assurance
- Institute of Statisticians
- Ministry of Defence
- Royal Statistical Society

This British Standard, having been prepared under the direction of the Quality, Management and Statistics Standards Policy Committee, was published under the authority of the Standards Board and comes into effect on 20 December 1991

© BSI 07-1999

First published as BS 2846, April 1957
Second edition as BS 2846-1, June 1975
Third edition December 1991

The following BSI references relate to the work on this standard:
Committee reference QMS/12
Draft for comment 86/60257 DC

ISBN 0 580 19793 X

Amendments issued since publication

Amd. No.	Date	Comments

Contents

	Page
Committees responsible	Inside front cover
Foreword	iii
Section 1. General principles	
0 Introduction	1
1 Scope	1
2 Definitions	1
3 Frequency tables	1
4 Graphical methods	1
5 Representative measures	1
6 Uncertainty of estimates	1
7 Rounding	2
8 Comparing estimates	2
9 General recommendations	3
Section 2. Rules for presentation of experimental results	
10 General	4
11 Example of presentation	4
Section 3. Rounding of observations	
12 Reasons for rounding	5
13 The effect of rounding on variation	5
14 Rounding interval: definition and choice	5
15 Caution on rounding	6
Section 4. Frequency distributions	
16 Purpose	7
17 Grouping into classes	7
18 Rules for grouping	7
19 Example of grouping	7
20 Presentation of tables	9
21 Graphical representation	12
22 Symmetry of distribution	15
Section 5. Further condensation: arithmetic mean and standard deviation	
23 Necessity for further condensation	16
24 Measures of central tendency	16
25 Measures of dispersion	16
26 Computation of the mean and of the standard deviation	17
27 Rounding of means and standard deviations	21
Section 6. Uncertainty in estimation	
28 General	22
29 Distributions of means and standard deviations	22
30 Confidence limits and confidence intervals	22
31 Homogeneity	25
32 The assumption of normality	27
Appendix A Presentation of numerical values (fineness of expression; rounding of numbers)	28
Figure 1 — Frequency distribution of compressive strengths of 400 concrete cubes (equal class widths)	13
Figure 2 — Frequency distribution of compressive strengths of 400 concrete cubes (general case)	13

	Page
Figure 3 — Frequency curve for compressive strengths of 400 concrete cubes	14
Figure 4 — Percentage cumulative distribution of compressive strengths of 400 concrete cubes	15
Figure 5 — Scatter diagram for % P_2O_5 in fertilizer (from Table 6)	15
Figure 6 — Control charts for compressive strength of 400 concrete cubes	26
Table 1 — Compressive strengths of concrete cubes (after 7 days)	8
Table 2 — Methods of classifying observations	10
Table 3 — Two ways of presenting a frequency distribution	11
Table 4 — Alternative way of defining classes in a frequency distribution	11
Table 5 — Two ways of presenting a cumulative frequency distribution	12
Table 6 — Analysis of fertilizer (% P_2O_5)	18
Table 7 — Analysis of fertilizer (% P_2O_5): (difference from 11.00 %)	19
Table 8 — Calculation of mean and standard deviation of a frequency distribution	21
Table 9 — Factors for control lines for mean and range charts	25
Publication(s) referred to	Inside back cover

Foreword

This Part of BS 2846 has been prepared under the direction of the Quality, Management and Statistics Standards Policy Committee. It supersedes BS 2846-1:1975 which is withdrawn.

The principal change made from the previous edition is the addition of an appendix on rounding of observations and estimates to augment the guidance on rounding already given. This new appendix, which is a shortened version of BS 1957:1953, replaces the previous Appendix A which explained the statistical terms used and is now no longer needed following the publication of BS 5532 "Statistical terminology". Other changes include the addition of text in clause 6 to introduce the concept of hypothesis testing and modifications of the techniques to take account of the use of calculators and computers. In particular, the guidance formerly given on the estimation of the standard deviation from the range has been deleted, since the standard deviation can be easily obtained directly from a calculator, many of which are pre-programmed for this purpose. (In practice, however, it seems likely that the mean range will continue to be used to calculate the numerical values for the action and warning lines used on control charts for the mean and the range.)

The importance of correct interpretation and presentation of test results has been increasingly recognized in the analysis of data obtained from manufacturing processes based on sample determinations and prototype evaluations in industry, commerce and educational institutions. It was for this reason that subcommittee 2 of Technical Committee 69 "Application of statistical methods" of the International Organization for Standardization (ISO), was charged with the task of establishing a guide to statistical methods for the interpretation of test results. International agreement has already been reached on the specification of several techniques and these have been published as Parts of BS 2846, as follows.

- *Part 1: Routine analysis of quantitative data;*
- *Part 2: Estimation of the mean: confidence interval (identical to ISO 2602);*
- *Part 3: Determination of a statistical tolerance interval (identical to ISO 3207);*
- *Part 4: Techniques of estimation and tests relating to means and variances (identical to ISO 2854);*
- *Part 5: Power of tests relating to means and variances (identical to ISO 3494);*
- *Part 6: Comparison of two means in the case of paired observations (identical to ISO 3301);*
- *Part 7: Tests for departure from normality (technically equivalent to ISO/DIS 5479).*

Part 1 is concerned with the routine analysis, presentation and reduction of quantitative data, such as rounding of test results, which should be considered prior to applying the statistical tests given in later Parts.

Part 2 specifies techniques for the calculation of confidence intervals for the mean of a normal population based on the results of a test applied to each of a random sample of individuals drawn from this population. It deals only with the case where the variance of the population is unknown.

Part 3 specifies techniques for the estimation of statistical tolerance limits, i.e. intervals containing, with a fixed probability, at least a given fraction of the population. It also includes a brief introduction to distribution-free methods of determining statistical tolerance limits, which are consequently valid when no assumptions can be made regarding the distribution of the underlying population.

Part 4 relates the estimates of the means and variances of a sample to the population parameters and specifies the techniques required to examine certain hypotheses concerning the numerical values of those parameters.

Part 5 on the other hand specifies procedures for assessing the risks of drawing wrong conclusions from sample data when applying the techniques of Part 4. It permits the determination of appropriate sample sizes to keep the risks down to acceptable levels, and the determination of the most appropriate acceptance criterion to keep the risks in balance.

Part 6 deals with the special case known as the method of paired comparisons, which Part 4 touches on briefly. Such cases can arise, for example, when two tests are performed on each specimen in the sample, a first prior to the application of some treatment and the second after, in order to ascertain whether the treatment is effective or not.

It is important to be aware of the various assumptions underlying the application of the methods, as stated in the text of Parts 1 to 6. These assumptions can be satisfied in almost any situation if the experiment has been designed properly from a statistical standpoint and this standard is concerned with data arising from experiments that are satisfactory in this respect. Although not dealt with in the following pages, the design of experiments is an important subject to which experimenters should give careful thought, more thought than is sometimes given before an experiment is started.

Most of the techniques for the analysis of quantitative data described in this standard assume that the items in the sample come from a population which has a normal distribution. In certain cases this assumption is not critical but in others, unless it can be shown to be at least approximately true, the techniques should not be applied. This applies particularly to the estimation of statistical tolerance intervals as specified in Part 3 and to the tests for variances in Part 4. This problem is discussed briefly in Parts 1 and 4 and in more detail in Part 7. Part 7 describes procedures which may be used to test whether apparent departures from normality can be regarded as sampling fluctuations or represent a significant degree of non-normality.

Standards giving guidance on other techniques being developed at international level are:

- Estimation of a median
- Estimation of a proportion
- Comparison of a sample proportion with a given value
- Comparison of two proportions

When these guides are published as international standards they will be considered for publication as dual-numbered British Standards, to provide further Parts of this standard.

This standard does not contain methods for the collection of individuals in the random sample. In this respect references should be made to the particular British Standards which specify the sampling procedures for a specific product. In the case of discrete items BS 6000, BS 6001 and BS 6002 contain general descriptions.

The following British Standards also provide practical guidance on the application of statistical methods: BS 600, BS 2564, BS 5497, BS 5700, BS 5701, BS 5702¹⁾, BS 5703, BS 6000, BS 6001²⁾, BS 6002.

¹⁾ In preparation.

²⁾ BS 6001-0 is in preparation.

A British Standard does not purport to include all the necessary provisions of a contract. Users of British Standards are responsible for their correct application.

Compliance with a British Standard does not of itself confer immunity from legal obligations.

Summary of pages

This document comprises a front cover, an inside front cover, pages i to vi, pages 1 to 32, an inside back cover and a back cover.

This standard has been updated (see copyright date) and may have had amendments incorporated. This will be indicated in the amendment table on the inside front cover.

Section 1. General principles

0 Introduction

It is generally agreed that reports of the results of investigations should preferably include the observations that have been made, unless limitations of space preclude this. Even when it is possible to give the observations in full, it is preferable also to give the essential information in a condensed form for ease of interpretation. It is desirable, not only to do this in a uniform way to avoid confusion, but also to ensure that the form adopted should follow the methods that have been found most satisfactory in practice.

1 Scope

This Part of BS 2846 gives rules which are recommended for use in the condensation, interpretation and presentation of data; it is, however, restricted to the consideration of observations of one variable only.

This Part also deals with the uncertainty of various quantities estimated from the data, arising from the complex of variations which are termed errors. The applicability of the various measures of uncertainty depends on the validity of certain assumptions, which are referred to in the text. Many of these assumptions are satisfied only if the experimental part of the investigation is designed properly from a statistical standpoint.

NOTE The titles of the publications referred to in this standard are listed on the inside back cover.

2 Definitions

For the purposes of this British Standard the definitions given in BS 5532 apply.

3 Frequency tables

Although it may be desirable to present the whole of the original data, it is not always convenient to do so if the data are numerous. While a few values may be reported in the text and cause no confusion, such a procedure with a large body of data might cause considerable interruption of the discussion. It is often better, therefore, to give full results in an appendix, and to include a condensed form of the data in the text. Some condensation is achieved by grouping. Grouped results may be presented in two ways, namely, as frequency tables or as cumulative frequency tables, and both ways are described in section 4. For most purposes grouping is of little value if fewer than thirty results are available.

4 Graphical methods

In general, it is not difficult to grasp the information in a comparatively small number (ten or fewer) of values. It is, however, beyond the ability of most people to visualize the general features of a mass of data, even when grouped into classes, and the use of diagrams will usually facilitate interpretation. The diagrams most frequently used for this purpose are stem-and-leaf displays, the histogram and the cumulative diagram; these are described in section 4.

5 Representative measures

Even the grouped data cannot be used readily for comparison with other data or information; further condensation is required for this purpose. Two statistics, the sample arithmetic mean and the sample standard deviation, are for many purposes adequate to represent the given data; these are described in section 5 and methods of calculating them are given. In some circumstances other statistics may be more appropriate. Some of these alternative statistics are also described in section 5.

6 Uncertainty of estimates

It is a matter of experience that observations made under apparently identical conditions usually show some variation, because of factors outside the experimenter's control. Such variation in the observations leads to related, but smaller, variation in means and in standard deviations calculated from the observations. Thus, since experiments when repeated usually give results with slightly different means and standard deviations, the mean and the standard deviation estimated from a single group of results are subject to some uncertainty.

This uncertainty can lead to results which apparently run counter to an established hypothesis about the process producing the data. It is then necessary to test whether this departure can reasonably be explained as due to chance fluctuation in the sampling results or whether a new hypothesis should be considered. The test procedure is formalized by calling the established hypothesis (expressed in statistical terms) the *null hypothesis* H_0 , and the other hypothesis an *alternative hypothesis* H_1 . The hypothesis H_0 might be that the data were normally distributed with mean 100 and variance 10, the alternative H_1 could be that the process mean was greater than 100 but the process variance and the distributional form were unchanged.

All possible results from a complete sample are then considered and classified according to the extent to which they favour the alternative hypothesis H_1 as the more plausible of the two hypotheses. A set of those results which most favour H_1 are then selected to form a critical region within which the null hypothesis H_0 will be rejected. The set of results is chosen so that there is a predetermined low probability of falling in the critical region if the null hypothesis is true. This is the probability of wrongly rejecting H_0 when it is in fact true. In general, many different sets of results could be chosen to form the critical region, so if it exists we choose that set for which the probability of falling in the critical region is always greatest when the alternative hypothesis H_1 is true.

The conclusions from the test are one of the following.

- a) The null hypothesis is rejected, (the result fell in the critical region).
- b) The null hypothesis is not rejected, (the data are inconclusive).
- c) The null hypothesis is not rejected, (the data suggest that the null hypothesis is the correct explanation).

The distinction between b) and c) is the extent of the uncertainty surrounding the estimated value.

Section 6 discusses this uncertainty and ways in which it may be represented, and Parts 2, 3 and 4 of this standard specify techniques for dealing with particular problems.

7 Rounding

It is frequently found that observations are reported to more figures³⁾ than are required for the purpose in view; this may be adjusted by rounding (see Appendix A).

Two forms of reporting are in common use.

- a) A fixed number of digits, called the number of significant figures, is recorded counting from the leftmost non-zero digit towards the rightmost significant digit. See A.3 a) for the ambiguous case when this digit is zero.
- b) A fixed number of decimal places is retained. This fixed number is positive if there are digits to the right of the decimal point and negative if there are one or more zeros immediately to the left of the decimal point for all recorded values. In this case a change of unit can simplify the presentation.

Unless there is a very wide range of values in repeated observations it is recommended that experimental values be expressed to a fixed number of decimal places. The subject is discussed in section 3, where a rule is given to determine the maximum degree of rounding which may be carried out. Caution needs to be observed, however, in rounding data, since it may later become desirable to make comparisons which require the greater number of figures originally available.

8 Comparing estimates

It should be borne in mind that this Part of BS 2846 gives rules only for the condensation and presentation of experimental results. While the purpose of such condensation is usually to permit comparisons to be made, the conditions which need to be fulfilled before such comparisons are valid are not discussed here, and reference should be made to statistical texts on the design and analysis of experiments. It may be noted, however, that any comparison made should be judged against an estimate of variability, e.g. a standard deviation, which takes account of all sources of error which might affect the comparison. The subject is discussed, in rather more detail, in clause 31 and Parts 4, 5 and 6 of this standard.

³⁾ In this standard the word "figure" is used as if synonymous with the word "digit". Zero is regarded as a figure except when expressly excluded.

9 General recommendations

This Part of BS 2846 includes a list of general recommendations, which for convenience of reference is given in section 2.

It should be stressed that, although the following sections give some explanation of the reasons for the adoption of the procedures recommended, this explanation is necessarily very brief.

Section 2. Rules for presentation of experimental results

10 General

10.1 Report experimental observations in full, unless limitations of space preclude this; if the observations are numerous they may be given in an appendix to avoid interrupting the discussion. Do not omit observations because they are outliers, or appear to be inconsistent with the rest of the data. State the units of measurement.

10.2 Present the results also as a frequency distribution, unless there are fewer than thirty observations. To facilitate this:

- a) rules or grouping are given in clause 18;
- b) a method of classification is described in clause 19;
- c) presentation of tables is discussed in clause 20.

If there are fewer than thirty observations, show their positions in a one-dimensional scatter diagram (see 21.4).

10.3 Give the following information:

- a) either the arithmetic mean (see 24.1, clause 26 and BS 2846-2), or the median if more appropriate (see 24.2, clause 26 and BS 2846-8⁴⁾;
- b) either the standard deviation (see 25.1, clause 26 and BS 2846-4); or the semi-interquartile range if more appropriate (see 25.2 and clause 26);
- c) the number observations.

NOTE This information should always be provided whether or not a frequency table is given.

10.4 If possible also give the following:

- a) 95 % confidence limits for the population mean (see clause 30 and BS 2846-2);
- b) 95 % confidence limits for the population standard deviation (see clause 30 and BS 2846-4).

10.5 Give all possible information on the conditions under which the observations were made. If the conditions were not constant for the whole set of observations, it is possible that the observations do not form a homogeneous group; homogeneity is discussed in clause 31. Where there is evidence of heterogeneity the data should be subdivided to give homogeneous subgroups; the conditions ruling for each subgroup should be stated. The mean, the standard deviation, and the number of observations should be given separately for each subgroup.

10.6 Examine whether observations could with safety be rounded (see section 3). If there is any likelihood that the data will be used later for purposes other than those intended originally, however, it is best not to round them.

11 Example of presentation

In section 4 the data for compressive strengths of concrete cubes given in Table 1 are shown as a frequency table (see Table 3) and as a histogram (see Figure 1).

When further condensed these results would be presented as follows.

		95 % confidence limits	
		Lower	Upper
Number of observations	400		
Arithmetic mean (N/mm ²)	27.36	27.01	27.71
Standard deviation (N/mm ²)	3.57	3.32	3.86

Some prefer, however, to attach the confidence limits directly to the relevant estimates. This might be done quite conveniently for the mean if the limits were given in brackets as shown below, and would be satisfactory provided a note was added which explained the meaning to be given to the figures in brackets.

Number of observations	Arithmetic mean
400	27.36 (± 0.35) ^a N/mm ²

^a The figures in brackets represent the 95 % confidence limits.

Unfortunately, the confidence limits for the *standard deviation* are not usually symmetrical about the estimate, so they cannot be expressed conveniently in this way.

⁴⁾ In preparation.

Section 3. Rounding of observations

12 Reasons for rounding

When data are obtained experimentally it is usual to record as many figures as possible without considering whether the figures retained are justified. In practice, results may appear more precise than they really are, or may be more precise than is required; it may be, therefore, that some improvement in data integrity could be achieved by omitting some of the figures recorded, without causing any serious loss of precision. The process of discarding any figures beyond the number of significant figures which it is desired to retain is known as rounding (sometimes called rounding-off). When this process is applied the last figure (or figures) retained is (are) adjusted if considered necessary. To ensure that large errors are not introduced it is necessary to use caution in deciding the extent to which rounding is to be applied.

NOTE Rounding is not the same thing as truncation, which is the process of simply discarding figures without any consideration of the last figure or figures retained.

Before proceeding with rounding of values the reader should consult Appendix A.

13 The effect of rounding on variation

All observations are subject to some variation. For example, two independent chemical analyses carried out on the same sample will rarely agree exactly because of minor variations in technique, dimensions of glassware, etc. From a series of *independent* duplicate results an estimate can be made of the variation arising from analysis alone. If the results of analysis are rounded (say) to one figure less it will usually be found that the apparent analytical variation has been increased; if the results are rounded to two figures less the variation will be further increased. This is because some variation arising from the process of rounding has been added. It is necessary therefore to set an upper limit to this additional variation and from this prescribed maximum to estimate the degree of rounding which is not to be exceeded.

14 Rounding interval: definition and choice

14.1 Definition

The amount of rounding which will be considered acceptable is specified by the *rounding interval*. This is defined as the smallest possible difference (excluding zero) between two rounded values in the same series.

14.2 Choice of rounding interval

The rounding interval chosen should be small enough for the additional variation arising from rounding to be small compared with the original variation in the observations. It can be shown that if the rounding interval is chosen to be equal to half the standard deviation of the observations the effect of rounding will be, on the average, to increase the standard deviation by only about 1 %. It will usually be found in practice that half the standard deviation is not a convenient unit to use as the rounding interval. Thus the rule is to choose the most convenient interval less than half the standard deviation. The adoption of this rule will ensure, on the average, that the rounding of observations will not increase the estimated standard deviation by more than 1 % in fact the increase will frequently be much less than this.

Attention is drawn to clause 15 and especially 15.2.

14.3 Examples of rounding

In 26.3, standard deviations are calculated for three sets of experimental results; these may be used to decide the extent to which the observations may be rounded in each of the following examples.

a) Fertilizer samples

$$s = 0.164 \% \text{ P}_2\text{O}_5;$$

$$s/2 = 0.082 \% \text{ P}_2\text{O}_5.$$

Using the rule in 14.2, therefore, the observations (see 26.3.1) for the fertilizer samples may be rounded to 0.05 % or 0.01 % P_2O_5 , whichever is considered the more convenient. In this example it is felt that the unit 0.01 % P_2O_5 is the better and that further rounding is inadvisable on grounds dealt with in clause 15.

b) Breaking strengths of hand drawn copper wire

$$s = 38 \text{ N};$$

$$s/2 = 19 \text{ N}.$$