ERRATA

28

International Standard ISO 4259

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

Petroleum products — Determination and application of precision data in relation to methods of test

Produits pétroliers — Détermination et application des valeurs de fidélité relatives aux méthodes d'essai

First edition — 1979-11-15

UDC 665.7 : 519.23 : 620.1

Ref. No. ISO 4259-1979 (E)

Descriptors : petroleum products, test results, reliability, statistical analysis, definitions

Price based on 27 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 4259 was developed by Technical Committee ISO/TC 28, *Petroleum products and lubricants,* and was circulated to the member bodies in April 1977.

It has been approved by the member bodies of the following countries :

Australia	Hungary	Romania
Austria	India	South Africa, Rep. of
Brazil	Iran	Spain
Bulgaria	Israel	Turkey
Canada	Italy	United Kingdom
Finland	Japan	USA
France	Poland	USSR
Germany, F.R.	Portugal	

The member body of the following country expressed disapproval of the document on technical grounds :

Sweden

Printed in Switzerland

Published 1981-01-15

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

Petroleum products — Determination and application of precision data in relation to methods of test

ERRATUM

Page 1

In clause 1, paragraph 2, line 2, replace "tests" by "test".

Page 4

In sub-clause 4.1, paragraph 3, line 3, replace "weighed" by "weighted".

Page 6

In sub-clause 4.3, paragraph 2, line 3, replace "If" by "It".

Page 7

In sub-clause 5.1.2, formula (4), replace "SS_1" by "$S'S_1$".

In sub-clause 5.2, line 4, add "at" after "out".

Page 17

In clause C.3, in the expression for "W_j^2", replace "$\sum_{j=1}^{S}$" by "$\sum_{i=1}^{L}$".

CONTENTS

Petroleum products — Determination and application of precision data in relation to methods of test

0 INTRODUCTION

For purposes of quality control and to check compliance with specifications, the properties of commercial petroleum products are assessed by standard laboratory test methods. Two or more measurements of the same property of a specific sample by any given test method do not usually give exactly the same result. It is therefore necessary to take proper account of this fact, by arriving at statistically based estimates of the precision for a method, i.e. an objective measure of the degree of agreement to be expected between two or more results obtained in specified circumstances.

1 SCOPE AND FIELD OF APPLICATION

This International Standard covers the calculation of precision estimates and their application to specifications. In particular, it contains definitions of relevant statistical terms (clause 2), the procedures to be adopted in the planning of an inter-laboratory test programme to determine the precision of a test method (clause 3), the method of calculating the precision from the results of such a programme (clauses 4 and 5), and the procedure to be followed in the interpretation of laboratory results in relation both to precision of the methods and to the limits laid down in specifications (clauses 6 to 9).

It must be emphasised that the procedures in this International Standard are designed to cover methods of tests for petroleum products only. The latter are, in general, homogeneous products with which serious sampling problems do not normally arise. It would not be appropriate, therefore, to consider the procedures to be necessarily of wider application, for example to heterogeneous solids.

2 DEFINITIONS

For the purposes of this International Standard, the following definitions apply :

2.1 analysis of variance : A technique which enables the total variance of a method to be broken down into its component factors.

2.2 between-laboratory variance : When results obtained by more than one laboratory are compared, the scatter is usually wider than when the same number of tests are carried out by a single laboratory, and there is some variation between means obtained by different laboratories. These give rise to the between-laboratory variance which is that component of the overall variance due to the difference in the mean values obtained by different laboratories. (There is a corresponding definition for between-operator variance.)

2.3 bias : The difference between the true value (related to the method of test) (see 2.24) and the known value (see 2.8), where this is available.

2.4 blind coding : The assignment of a different number to each sample but not to repeats. No other identification or information on any sample is given to the operator.

2.5 check sample : A sample taken at the place where the product is exchanged, i.e. where the responsibility for the product quality passes from the supplier to the recipient.

2.6 degrees of freedom : The divisor used in the calculation of variance; one less than the number of independent results.

NOTE — The definition applies strictly only in the simplest cases. Complete definitions are beyond the scope of this International Standard.

2.7 determination : The process of carrying out the series of operations specified in the test method, whereby a single value is obtained.

2.8 known value : The actual quantitative value implied by the preparation of the sample.

NOTE — The known value does not always exist, for example for empirical tests such as flash point.

2.9 mean; arithmetic mean; average : For a given set of results, the sum of the results divided by their number.

2.10 mean square : The sum of squares divided by the degrees of freedom.

2.11 normal distribution : The probability distribution of a continuous random variable X such that, if x is any real number, the probability density is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad \ldots (1)$$

$$-\infty < x < +\infty$$

NOTE – μ is the true value and σ is the standard deviation of the normal distribution ($\sigma > 0$).

2.12 operator : A person who normally and regularly carries out a particular test.

2.13 outlier : A result far enough in magnitude from other results to be considered not a part of the set.

2.14 precision : The closeness of agreement between the results obtained by applying the experimental procedure several times on identical materials and under prescribed conditions. The smaller the random part of the experimental error, the more precise is the procedure.

2.15 random error : The chance variation encountered in all test work despite the closest control of variables.

2.16 recipient : Any individual or organization who receives or accepts the product delivered by the supplier.

2.17 repeatability :

a) *Qualitatively*

The closeness of agreement between successive results obtained in the normal and correct operation of the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory, and short intervals of time).

NOTE – The representative parameters of the dispersion of the population which may be associated with the results are qualified by the term "repeatability", for example repeatability standard deviation, repeatability variance.

b) *Quantitatively*

The value equal to or below which the absolute difference between two single test results obtained in the above conditions may be expected to lie with a specified probability; in the absence of other indication, the probability level is 95 %.

2.18 replication : The execution of a test method more than once so as to improve precision and to obtain a closer estimation of sampling error. Replication should be distinguished from repetition in that the former implies that experiments are carried out at one place and, as far as possible, one period of time.

2.19 reproducibility :

a) *Qualitatively*

The closeness of agreement between individual results obtained in the normal and correct operations of the same method on identical test material but under different conditions (different operators, different apparatus and different laboratories).

NOTE – The representative parameters of the dispersion of the population which may be associated with the results are qualified by the term "reproducibility", for example reproducibility standard deviation, reproducibility variance.

b) *Quantitatively*

The value equal to or below which the absolute difference between two single test results on identical material obtained by operators in different laboratories, using the standardized test method, may be expected to lie with a specified probability; in the absence of other indication, the probability level is 95 %.

2.20 result : The final value obtained by following the complete set of instructions in the test method; it may be obtained from a single determination of from several determinations depending on the instructions in the method. (It is assumed that all the results are rounded off according to the procedure specified in annex G.)

2.21 standard deviation : A measure of the dispersion of a series of results around their mean, equal to the positive square root of the variance and estimated by the positive square root of the mean square.

2.22 sum of squares : The sum of squares of the differences between a series of results and their mean.

2.23 supplier : Any individual or organization responsible for the quality of a product just before it is taken over by the recipient.

2.24 true value : For practical purposes, the value towards which the average of single results obtained by n laboratories tends, as n tends towards infinity; consequently, such a true value is associated with the particular method of test.

NOTE – A different and idealized definition is given in ISO 3534, *Statistics – Vocabulary and symbols.*

2.25 variance : The mean of the squares of the deviation of a random variable from its mean.

3 STAGES IN PLANNING OF AN INTER-LABORATORY TEST PROGRAMME FOR THE DETERMINATION OF THE PRECISION OF A TEST METHOD

The stages in planning an inter-laboratory test programme are as follows :

a) Preparing a draft method of test.

b) Planning a pilot programme with two laboratories.

c) Planning the inter-laboratory programme.

d) Executing the inter-laboratory programme.

The four stages are described in turn.

3.1 Preparing a draft method of test

This shall contain all the necessary details for carrying out the test and reporting the results. Any condition which could alter the results shall be specified.

The clause on precision will be included at this stage only as a heading.

3.2 Planning a pilot programme with at least two laboratories

A pilot programme is necessary for the following reasons :

a) to verify the details in the operation of the test;

b) to find out how well operators can follow the instructions of the method;

c) to check the precautions regarding samples;

d) to estimate roughly the precision of the test.

At least two samples are required, covering the range of results to which the test is intended to apply; however, at least 12 laboratory/sample combinations should be included. Each sample is tested twice by each laboratory under repeatability conditions. If any omissions or inaccuracies in the draft method are revealed, they shall now be corrected. The results shall be analysed for bias and precision : if either is considered to be too large, then alterations to the method shall be considered.

3.3 Planning the inter-laboratory programme

There shall be at least five participating laboratories, but it is preferable to exceed this number in order to reduce the number of samples required.

The number of samples shall be sufficient to cover the range of the property measured, and to give reliability to the precision estimates. If any variation of precision with level was observed in the results of the pilot programme, then at least five samples shall be used in the inter-laboratory programme. In any case, it is advisable to aim for 30 degrees of freedom in both repeatability and reproducibility. For repeatability, this means obtaining a total of 30 pairs of results in the programme. For reproducibility, table 11 (annex A) gives the number of samples required in terms of L, P and Q, where L is the number of participating laboratories and P and Q are the ratios of variance component estimates obtained from the pilot programme. Specifically, P is the ratio of the interaction component to the repeats component, and Q is the ratio of the laboratories component to the repeats component. Annex B gives the derivation of the formula used. If Q is much larger than P, then 30 degrees of freedom cannot be achieved; the blank entries in table 11 correspond to this situation or the approach of it (i.e. when more than 20 samples are required). For these cases, there is likely to be a significant bias between laboratories.

3.4 Executing the inter-laboratory programme

One person shall be responsible for the entire programme, from the distribution of the texts and samples, to the final appraisal of the results. He shall be familiar with the method, but shall not personally take part in the tests.

The text of the method shall be distributed to all the laboratories in time to raise any queries before the tests begin. If any laboratory wants to practice the method in advance, this shall be done with samples other than those used in the programme.

The samples shall be accumulated, subdivided and distributed by the organizer, who shall also keep a reserve of each sample for emergencies. It is most important that the individual laboratory portions be homogeneous. They shall be blind-coded before distribution, and the following instructions shall be sent with them :

a) the agreed draft method of test;

b) the handling and storing requirements for the samples;

c) the order in which the samples are to be tested (a different random order for each laboratory);

d) the statement that two results are to be obtained consecutively on each sample by the same operator with the same apparatus;

e) the period of time during which all the samples are to be tested;

f) a form for reporting the results. For each sample, there shall be space for the date of testing, the two results, and any unusual occurrences. The unit of accuracy for reporting the results shall be specified;

g) a statement that the test shall be carried out under normal conditions, using operators with good experience but not exceptional knowledge; and that the duration of the test shall be the same as normal.

NOTE — The pilot programme operators may take part in the inter-laboratory programme. If their extra experience in testing a few more samples produces a noticeable effect, it should serve as a warning that the method is not satisfactory. They should be identified in the report of the results so that any effect may be noted.

4 INSPECTION OF INTER-LABORATORY RESULTS FOR UNIFORMITY AND FOR OUTLIERS

4.0 Introduction

This clause specifies procedures for examining the results reported in a statistically designed inter-laboratory programme (see clause 3) to establish

a) the independence of precision,

b) the level of the results,

c) the uniformity of precision from laboratory to laboratory,

and to detect the presence of outliers. The procedures are described in mathematical terms based on the notation of annex C and illustrated with reference to the example of calculation of the bromine number set out in annex D.

Throughout this clause (and clause 5), the procedures to be used are first specified and then illustrated by a worked example using data given in annex D.

It is assumed throughout this clause that all the results are either from a single normal distribution or capable of being transformed into such a distribution (see 4.1). Other cases (which are rare) would require different treatment which is beyond the scope of this International Standard.

4.1 Transformation of data

In many test methods the precision depends on the level of the test result, and thus the variability of the reported results is different from sample to sample. The method of analysis outlined in this International Standard requires that this should not be so and the position is rectified, if necessary, by a transformation.

The laboratories standard deviations D_j, (see annex C, clause C.3) are calculated and plotted against the sample means m_j. If the points so plotted may be considered as lying about a line parallel to the m-axis, then no transformation is necessary. If, however, the plotted points lie about a curve of the form $D = f(m)$, then a transformation will be necessary.

The relationship $D = f(m)$ is best established by the technique of univariate regression analysis (strictly speaking, an iteratively weighed regression should be used, but in most cases an unweighted regression gives a satisfactory approximation).

An outline of the calculation necessary is given in annex F, but it is a standard programme for most computers. Normally, a 5 % significance level will be used to test whether a regression coefficient differs from zero.

If it has been shown that there is a significantly non-zero regression coefficient giving a dependence of the form $D = f(m)$, then the appropriate transformation $y = F(x)$, where x is the reported result, is given by the formula

$$F(x) = k \int \frac{dx}{f(x)} \qquad \ldots (2)$$

where k is constant.

The particular cases likely to be encountered, together with the required transformations, are listed in table 20 (annex E). A regression of log D_j on log m_j will show any dependence of the form $D = Am^B$.

The choice of transformation is difficult to make the subject of formalized rules and qualified statistical assistance may be required in particular cases.

After selecting a transformation on the basis of the dependence of D on m, it shall be verified that the same transformation is also relevant for the repeats standard deviation d (see annex C, clause C.3). If it is not, then either a separate transformation will be necessary or the results will not need transforming for the calculation of repeatability.

4.1.1 *Worked example*

Table 1 lists the values of m, D, and d for the eight samples in the example given in annex D.

Inspection of the figures in table 1 shows that both D and d increase with m, the rate of increase diminishing as m increases. A plot of these figures on log-log paper (i.e. a graph of log D and log d against log m) shows that the points may reasonably be considered as lying about two straight lines (see the figure in annex F). The gradients of these lines are 0,64 and 0,58 respectively and thus, bearing in mind the errors in these estimated gradients, they may for convenience be considered as parallel lines with gradient 2/3.

Hence, the same transformation is appropriate both for repeatability and reproducibility, and is given by the formula

$$\int x^{-2/3}\,dx = 3\,x^{1/3} \qquad \ldots (3)$$

Since the constant multiplier may be ignored, the transformation thus reduces to that of taking the cube roots of the reported results (bromine numbers). This yields the transformed data shown in table 16 (annex D), in which the cube roots are quoted correct to three decimal places.

4.2 Tests for outliers

After application of the appropriate transformation (or transformations) to the reported data, or if it has been decided that this is not necessary, the transformed results shall be inspected for outliers. These are the values which are so different from the remainder that it can only be concluded that they have arisen from some fault in the application of the method or from testing a wrong sample. Many possible tests may be used and the associated significance levels varied, but those that are specified in the following sub-clauses have been found to be appropriate in this International Standard.

4.2.1 *Uniformity of repeatability*

The first outlier test is concerned with detecting a discordant result in a pair of repeat results. This test[1] involves calculating the e_{ij}^2 over all the laboratory/sample

TABLE 1

Sample number	3	8	1	4	5	6	2	7
m	0,756	1,22	2,15	3,64	10,9	48,2	65,4	114
D	0,067	0,159	0,729	0,211	0,291	1,50	2,22	2,93
d	0,050 0	0,057 2	0,127	0,115	0,094 3	0,527	0,817	0,935

combinations. Cochran's criterion at the 1 % level is then used to test the ratio of the largest of these values over their sum (see annex C, clause C.4). If its value exceeds the value given in table 17 (annex D), corresponding to the 1 % probability level, k being the number of pairs available for comparison, then the member of the pair farthest from the sample mean shall be rejected and the process repeated, reducing k by 1, until no more rejections are called for. In certain cases, this test "snowballs" and leads to an unacceptably large proportion of rejections, (say more than 10 %). If this is so, this rejection test shall be abandoned and some or all of the rejected results shall be retained. An arbitrary decision based on judgement will be necessary in this case.

4.2.2 *Worked example*

In the case of the example given in annex D, the difference between transformed repeat results, i.e. of the pairs of numbers in table 16, in units of the third decimal place, are shown in table 2.

TABLE 2

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	42	21	7	13	7	10	8	0
B	23	12	12	0	7	9	7	0
C	0	6	0	0	7	8	4	0
D	14	6	0	13	0	8	9	32
E	65	4	0	0	14	5	7	28
F	23	20	34	29	20	30	43	0
G	62	4	78	0	0	16	18	56
H	44	20	29	44	0	27	4	32
J	0	59	0	40	0	30	26	0

The largest range is 0,078 for laboratory G on sample 3. The sum of squares of all the ranges is

$$0{,}042^2 + 0{,}021^2 + \ldots + 0{,}026^2 + 0^2 = 0{,}043\,9$$

Thus, the ratio to be compared with the Cochran's criterion is

$$\frac{0{,}078^2}{0{,}043\,9} = 0{,}138$$

There are 72 ranges and, as from table 17 (annex D), the criterion for 75 ranges is 0,180 9, this ratio is not significant.

4.2.3 *Uniformity of reproducibility*

The remaining outlier tests are concerned with establishing uniformity in the reproducibility estimate, and are designed to detect either a discordant pair of results from a laboratory on a particular sample or a discordant set of results from a laboratory on all samples. For both purposes, one of the range of Dixon r tests[2] is appropriate. This involves forming for each sample, and finally for the laboratory totals (see 5.2), ratios of various differences between the pair sums a_{ij} (see annex C, clause C.5).

The appropriate ratio shall be compared with the critical 1 % values given in table 18 (annex D), with the value of n determined by the number of laboratories concerned. If significant value is encountered for individual samples, the corresponding extreme values shall be omitted and the process repeated. If any extreme values are found in the laboratory totals, then all the results from this laboratory shall be rejected.

4.2.3.1 WORKED EXAMPLE

The application of Dixon's test to sample 1 is shown in detail below. (See note.)

The first step is to place the pair sums for each laboratory which tested sample 1 in ascending order of magnitude, as shown in table 3.

The appropriate Dixon ratio for nine laboratories is r_{11}.

For testing the highest value,

$$r_{11} = \frac{3{,}188 - 2{,}562}{3{,}188 - 2{,}409}$$

$$= 0{,}804$$

This value is greater than the tabulated value and so the results from laboratory D on this sample are rejected.

As there has been a rejection, the procedure is repeated for high values without the results for laboratory D being taken into account. This gives

$$r_{11} = \frac{2{,}562 - 2{,}540}{2{,}562 - 2{,}409}$$

$$= 0{,}144$$

Comparison of this value with the corresponding value in table 18 (annex D), for eight laboratories shows that it is not significant and so there are no further outliers.

For testing the lowest value,

$$r_{11} = \frac{2{,}409 - 2{,}409}{2{,}562 - 2{,}409}$$

$$= 0$$

This value is compared with the corresponding value in table 18 (annex D), namely 0,677.

TABLE 3

Laboratory	B	F	C	H	E	A	G	J	D
Pair sum	2,409	2,409	2,432	2,476	2,497	2,520	2,540	2,562	3,188

As the calculated value is less than the one in table 18, there are no outliers at the low end.

This procedure is repeated for each sample. In this example there were no further significant ratios, and so the only rejections made were those for sample 1 obtained by laboratory D.

NOTE – If the two lowest values or the two highest values are equal, there can be no corresponding outlier.

4.3 Rejection of complete data from a sample

The laboratories standard deviation and repeats standard deviation shall be examined for any outlying samples. If a transformation has been carried out or any rejection made, new standard deviations shall be calculated.

If the standard deviation for any sample is excessively large, it shall be examined with a view to rejecting the results from that sample. If is not possible to give an exact criterion for defining "excessively large" in this context, but it is felt that this action should be taken only in extreme cases[1].

NOTE – At this stage it is desirable to check that the rejections carried out have not invalidated the transformation used. If necessary, the procedure from 4.1 should be repeated with the outliers deleted.

4.3.1 *Worked example*

The laboratories standard deviations of the transformed results, after the rejection of the pair of results by laboratory D on sample 1, are given in table 4 in ascending order of sample mean.

Inspections shows that there is no outlying sample amongst these. It will be noted that the laboratories standard deviations are now independent of the sample means, which was the purpose of transforming the results. It was not considered necessary in this case to repeat the calculations with the outlier deleted.

The figures in table 5, taken from a test programme on bromine numbers over 100, will illustrate the case of a sample rejection.

It is clear, by inspection, that the laboratories standard deviation of sample 93 at 15,26 is far greater than the others, which are close together lying between 3,85 and 5,10, and so should be rejected. It is noted that the size of the repeats standard deviation in this sample also tends to confirm it as an outlier.

5 ANALYSIS OF VARIANCE AND CALCULATION OF PRECISION ESTIMATES

5.0 Introduction

After the data have been inspected for uniformity, a transformation has been performed if necessary, and any outliers have been rejected (see clause 4), an analysis shall be carried out. First the missing values shall be estimated by the least squares method, then an analysis of variance table constructed, and finally the precision estimates derived.

5.1 Estimating missing or rejected values

5.1.1 *One of the two repeat values missing or rejected*

If one of a pair of repeats (y_{ij1} or y_{ij2}) is missing or rejected, this shall be considered to have the same value as the other repeat in accordance with the least squares method.

5.1.2 *Both repeat values missing or rejected*

If both the repeat values are missing, estimates of a_{ij} ($= y_{ij1} + y_{ij2}$) shall be made by forming the laboratories × samples interaction sum of squares, including the missing values of the totals of the laboratories/samples pairs of results as unknown variables. Any laboratory from which all the results were rejected by Dixon's test shall be ignored and the new value of L used. The estimates of the missing or rejected values shall then be found by forming the partial derivatives of this sum of squares with respect to each variable in turn and equating these to zero to solve as a set of simultaneous equations.

Formula (4) may be used where only one pair sum has to be estimated. If more estimates are to be made, see, for instance, reference [5] for details.

TABLE 4

Sample number	3	8	1	4	5	6	2	7
Sample mean	0,910 1	1,066	1,240	1,538	2,217	3,639	4,028	4,851
Laboratories standard deviation	0,027 8	0,047 4	0,035 7	0,029 7	0,019 6	0,037 8	0,044 8	0,041 6

TABLE 5

Sample number	90	89	93	92	91	94	95	96
Sample mean	96,1	99,8	119,3	125,4	126,0	139,1	139,4	159,5
Laboratories standard deviation	5,10	4,20	15,26	4,40	4,09	4,87	4,74	3,85
Repeats standard deviation	1,13	0,99	2,97	0,91	0,73	1,32	1,12	1,36

1) A test which may prove to be appropriate, but of which no experience is available in this context, is that which involves the ratio of the maximum to the minimum of a set of variances (at the 1 % level), as described in *Biometrika tables for statisticians*, volume 1, table 31.

If the value of one pair sum a_{ij} has to be estimated, the estimate is given by the formula

$$a_{ij} = \frac{1}{(L-1)(S'-1)} (LL_1 + SS_1 - T_1) \quad \ldots (4)$$

where

$S' = S$ – number of samples rejected in 4.3

L_1 is the total of remaining pairs in the ith laboratory;

S_1 is the total of remaining pairs in the jth sample;

T_1 is the total of all pairs except a_{ij}.

5.1.2.1 WORKED EXAMPLE

The two results from laboratory D on sample 1 were rejected (see 4.2) and thus a_{41} has to be estimated.

Total of remaining results in laboratory 4 = 36,354

Total of remaining results in sample 1 = 19,845

Total of all the results except a_{41} = 348,358

Also $S' = 8$ and $L = 9$.

Hence, the estimate of a_{41} is given by

$$a_{ij} = \frac{1}{(9-1)(8-1)} [(9 \times 36{,}354) + (8 \times 19{,}845) - 348{,}358]$$

Therefore

$$a_{ij} = \frac{137{,}588}{56} = 2{,}457$$

5.2 Rejection test for outlying laboratories

At this stage, one further rejection test remains to be carried out. This determines whether it is necessary to reject the complete set of results from any particular laboratory. It could not be carried out an earlier stage, except in the case where no individual results or pairs are missing or rejected. The procedure again consists of the set of Dixon r tests (see 4.2.3), applied to the laboratory totals over all samples, with any estimated results included (see note in 4.3). If any laboratories are rejected on all samples, new estimates must be calculated for any remaining missing values (see 5.1).

5.2.1 *Worked example*

The procedure on the laboratory totals shown in table 6 below follows exactly that specified in 4.2.3.

Table 7 summarizes the calculated Dixon ratios.

TABLE 7

Ratio	r_{11}
Low	0,282
High	0,095

Comparison with the value tabulated in table 18 (annex D) shows that neither of these ratios is significant and therefore no complete laboratory rejections are necessary.

5.3 Analysis of variance

5.3.1 *Forming the sums of squares for the laboratories × samples interaction sum of squares*

The estimated values, if any, shall be put in the array and an approximate analysis of variance performed.

Mean correction

$$M = T^2/2L'S' \quad \ldots (5)$$

where $L' = L$ – number of laboratories rejected in 5.2.

Samples sum of squares

$$= \left[\sum_{j=1}^{S'} (g_j^2/2L') \right] - M \quad \ldots (6)$$

Laboratories sum of squares

$$= \left[\sum_{i=1}^{L'} (h_i^2/2S') \right] - M \quad \ldots (7)$$

Pairs sum of squares

$$= 1/2 \left[\sum_{i=1}^{L'} \sum_{j=1}^{S'} a_{ij}^2 \right] - M \quad \ldots (8)$$

I = the laboratories × samples interaction sum of squares

= (pairs sum of squares) – (laboratories sum of squares) – (sample sum of squares)

Ignoring any pairs in which there are estimated values,

E = repeats sum of squares

$$= 1/2 \sum_{i=1}^{L'} \sum_{j=1}^{S'} e_{ij}^2 \quad \ldots (9)$$

TABLE 6

Laboratory	G	C	H	D	A	B	E	F	J
Total	38,560	38,777	38,840	38,811*	38,992	39,020	39,099	39,329	39,387

* Including estimated value.

The purpose of performing this approximate analysis of variance is to obtain the minimized laboratories × samples interaction sum of squares, I. This is then used as indicated in 5.3.1.1, to obtain the laboratories sum of squares.

If there were no estimated values, the above analysis of variance is exact and the following sub-clause should be disregarded.

5.3.1.1 WORKED EXAMPLE

Mean correction

$$= \frac{350{,}815^2}{144}$$

$$= 854{,}660\ 5$$

Samples sum of squares

$$= \frac{22{,}302^2 + 72{,}512^2 + \ldots + 19{,}192^2}{18} - 854{,}660\ 5$$

$$= 293{,}540\ 9$$

Laboratories sum of squares

$$= \frac{38{,}992^2 + 39{,}020^2 + \ldots + 39{,}387^2}{16} - 854{,}660\ 5$$

$$= 0{,}035\ 6$$

Pairs sum of squares

$$= 1/2\ (2{,}520^2 + 8{,}041^2 + \ldots + 2{,}238^2) - 854{,}660\ 5$$

$$= 293{,}690\ 8$$

Repeats sum of squares

$$= 1/2\ (0{,}042^2 + 0{,}021^2 + \ldots + 0^2)$$

$$= 0{,}021\ 9$$

Table 8 can then be derived.

TABLE 8

Source of variation	Sum of squares
Samples	293,540 9
Laboratories	0,035 6
Laboratories × samples	0,114 3
Pairs	293,690 8
Repeats	0,021 9

5.3.2 *Forming the sum of squares for the exact analysis of variance*

In this sub-clause, all the estimated pairs are disregarded and new values of g are calculated. The following sums of squares for the exact analysis of variance[3] are formed.

Uncorrected sample sum of squares

$$= \sum_{j=1}^{S'} \frac{g_j^2}{S_j} \qquad \ldots (10)$$

Uncorrected pairs sum of squares

$$= 1/2 \sum_{i=1}^{L'} \sum_{j=1}^{S'} a_{ij}^2 \qquad \ldots (11)$$

where $S_j = 2$ (L' − number of missing pairs in that sample).

The laboratories sum of squares is equal to (pairs sum of squares) − (samples sum of squares) − (the minimized laboratories × samples interaction sum of squares).

$$1/2 \left[\sum_{i=1}^{L'} \sum_{j=1}^{S'} a_{ij}^2 \right] - \left[\sum_{j=1}^{S'} \frac{g_j^2}{S_j} \right] - 1 \qquad \ldots (12)$$

5.3.2.1 WORKED EXAMPLE

Uncorrected samples sum of squares

$$= \frac{19{,}845^2}{16} + \frac{72{,}512^2}{18} + \ldots + \frac{19{,}192^2}{18}$$

$$= 1\ 145{,}183\ 4$$

Uncorrected pairs sum of squares

$$= \frac{2{,}520^2}{2} + \frac{8{,}041^2}{2} + \ldots + \frac{2{,}238^2}{2}$$

$$= 1\ 145{,}332\ 9$$

Therefore, laboratories sum of squares

$$= 1\ 145{,}332\ 9 - 1\ 145{,}183\ 4 - 0{,}114\ 3$$

$$= 0{,}035\ 2$$

5.3.3 *Degrees of freedom*

The degrees of freedom for the laboratories are ($L' - 1$). The degrees of freedom for laboratories × samples interaction are ($L' - 1$) ($S' - 1$) for a complete array and are reduced by one for each pair which is estimated. The degrees of freedom for repeats are ($L'S'$) and are reduced by one for each pair in which one or both values are estimated.

5.3.3.1 WORKED EXAMPLE

There are eight samples and nine laboratories in this example. As no complete laboratories or samples were rejected, then $S' = 8$ and $L' = 9$.

Laboratories degrees of freedom = $L - 1 = 8$

Laboratories × samples interaction degrees of freedom, if there has been no estimates, would have been $(9 - 1)$ $(8 - 1) = 56$. But one pair was estimated, hence laboratories × samples interaction degrees of freedom = 55. Repeats degrees of freedom would have been 72 if there had been no estimates. In this case one pair was estimated, hence repeats degrees of freedom = 71.

5.3.4 *Mean squares and analysis of variance*

The mean square in each case is the sum of squares divided by the degrees of freedom. This leads to the analysis of variance shown in table 9.

TABLE 9

Source of variation	Degrees of freedom	Sum of squares	Mean square
Laboratories	$L'-1$	Laboratories sum of squares	M_L
Laboratories × samples	$(L'-1)(S'-1)$ – number of estimated pairs	I	M_{LS}
Repeats	$L'S'$ – number of pairs in which one or both values are estimated	E	M_r

5.3.4.1 WORKED EXAMPLE

The analysis of variance is shown in table 10.

TABLE 10

Source of variation	Degrees of freedom	Sum of squares	Mean square
Laboratories	8	0,035 2	0,004 400
Laboratories × samples	55	0,114 3	0,002 078
Repeats	71	0,021 9	0,000 308

5.4 Expectation of mean square and calculation of precision estimates

5.4.1 *Expectation of mean square with no estimated values*

For a complete array with no estimated values, the expectations of mean squares are

Laboratories : $\sigma_0^2 + 2\,\sigma_1^2 + 2S\,\sigma_2^2$

Laboratories × samples : $\sigma_0^2 + 2\,\sigma_1^2$

Repeats : σ_0^2

where

σ_1^2 is the component of variance due to interaction between laboratories and samples;

σ_2^2 is the component of variance due to differences between laboratories.

5.4.2 *Expectation of mean square with estimated values*

The coefficients of σ_1^2 and σ_2^2 in the expression for mean square are altered in the cases where there are estimated values. The expectations of mean squares then become

Laboratories : $\sigma_0^2 + \alpha\,\sigma_1^2 + \beta\,\sigma_2^2$

Laboratories × samples : $\sigma_0^2 + \gamma\,\sigma_1^2$

Repeats : σ_0^2

where

$$\alpha = \sum_{i=1}^{L'}\left(\sum_{j=1}^{S'} n_{ij}^2\right)\left(\frac{1}{N_i}-\frac{1}{N'}\right)\Bigg/(L'-1)$$

$$\beta = \left(N' - \frac{1}{N'}\sum_{i=1}^{L'} N_i^2\right)\Bigg/(L'-1)$$

$$\gamma = \left(N' - \frac{1}{N'}\sum_{i=1}^{L'}\sum_{j=1}^{S'} n_{ij}^2\right)\Bigg/(K-1)$$

where

n_{ij} is the number of results from the ith laboratory on the jth sample, if tested;

N_i is the number of results in ith laboratory;

N' is the total number of actual results minus number of rejected results;

K is the number of $L \times S$ cells containing a result.

If there are no cells with only a single estimated result, then

$$\alpha = \gamma = 2$$

NOTE – This sub-clause is based upon the assumptions that both samples and laboratories are "random effects".

5.4.2.1 WORKED EXAMPLE

For the example, which has eight samples, and nine laboratories.

$$\alpha = \left[7 \times 2^2\left(\frac{1}{14}-\frac{1}{142}\right) + 8 \times 8 \times 2^2\left(\frac{1}{16}-\frac{1}{142}\right)\right]\Bigg/8$$

$$= 2$$

$$\beta = \left[142 - \frac{(8 \times 16^2 + 14^2)}{142}\right]\Bigg/8$$

$$= 15{,}78$$

$$\gamma = \left[142 - \frac{(71 \times 2^2)}{142}\right]\Bigg/70$$

$$= 2$$

5.4.3 *Calculation of precision estimates*

5.4.3.1 REPEATABILITY

The repeatability variance is twice the mean square for repeats. The repeatability estimate is the product of the repeatability standard deviation and the "t-value" with appropriate degrees of freedom [see table 19 (annex D].

This calculated estimate shall be rounded down to the nearest unit used in quoting the results, as a consequence of the definition of repeatability.

5.4.3.1.1 *Worked example*

Repeatability variance

$= 2\,\sigma_0^2$

$= 0{,}000\,616$

Repeatability of y

$= t_{71}\sqrt{0{,}000\,616}$

$= 0{,}049\,5$

Repeatability of x

$= 3\,x^{2/3} \times 0{,}049\,5$

$= 0{,}148\,x^{2/3}$

5.4.3.2 REPRODUCIBILITY

Reproducibility variance $= 2\,(\sigma_0^2 + \sigma_1^2 + \sigma_2^2)$ and can be calculated using formula (13).

Reproducibility variance

$$= \frac{2}{\beta} M_L + \frac{2}{\gamma\beta}(\beta - \alpha)\, M_{LS} + \frac{2}{\gamma\beta}(\alpha - \beta - \gamma + \gamma\beta)\, M_r \quad \ldots (13)$$

where the symbols are as set out in 5.3.4 and 5.4.2.

The reproducibility estimate is the product of the reproducibility standard deviation and the "*t*-value" with appropriate degrees of freedom [see table 19 (annex D)]. An approximation to the degrees of freedom of the reproducibility variance is given by formula (14).

$$\nu = \frac{(\text{Reproducibility variance})^2}{\dfrac{r_1^2}{L' - 1} + \dfrac{r_2^2}{\nu_{LS}} + \dfrac{r_3^2}{\nu_r}} \quad \ldots (14)$$

where

r_1, r_2 and r_3 are the three successive terms in formula (13);

ν_{LS} is the degrees of freedom for laboratories × samples;

ν_r is the degrees of freedom for repeats.

This calculated estimate shall also be rounded down to the nearest unit used in quoting the results, as a consequence of the definition of reproducibility.

Note that if a transformation $y = f(x)$ has been used, then

$$R(x) \approx \left|\frac{dx}{dy}\right| R(y) \quad \ldots (15)$$

where $R(x)$, $R(y)$ are the corresponding reproducibility functions. A similar relationship applies to the repeatability functions $r(x)$, $r(y)$.

5.4.3.2.1 *Worked example*

Reproducibility variance

$$= \left(\frac{2}{15{,}78} \times 0{,}004\,40\right) + \left(\frac{13{,}78}{15{,}78} \times 0{,}002\,078\right) + 0{,}000\,308$$

$= 0{,}000\,558 + 0{,}001\,815 + 0{,}000\,308$

$= 0{,}002\,681$

$\nu = 7\,188/(39 + 60 + 1)$

$= \dfrac{7\,188}{100}$

$= 72$ (correct to nearest integer)

Reproducibility of y

$= t_{72}\sqrt{0{,}002\,681}$

$= 0{,}103\,4$

Reproducibility of x

$= 0{,}310\,x^{2/3}$

5.5 Precision clause of a method of test

When the precision of a method of test has been determined, according to the procedures set out in this International Standard, it shall be included in the method as follows :

"Range or sample description	Repeatability	Reproducibility
. . . . .	. . . . .	. . . . .
. . . . .	. . . . .	. . . . .

These precision values have been obtained using the procedures set out in ISO 4259."

6 SIGNIFICANCE OF REPEATABILITY *r* AND REPRODUCIBILITY *R* AS DISCUSSED IN EARLIER CLAUSES

6.0 Introduction

The value of these quantities is estimated from analysis of variance (two-factor with replication) performed on the results obtained in a statistically designed inter-laboratory programme in which different laboratories each test a range of samples. Repeatability and reproducibility values should be included in each published test method, and it is noted that the latter is usually greater than the former, if derived in accordance with this International Standard.

6.1 Repeatability *r*

Most laboratories do not carry out more than one test on each sample for routine quality control purposes, except in abnormal circumstances, such as in cases of dispute or if the test operator wishes to confirm that his technique is satisfactory. In these abnormal circumstances, when multiple results are obtained, it is useful to check the consistency of repeat results against the repeatability of the method and

the appropriate procedure is outlined in 6.1.1. It is also useful to know what degree of confidence can be placed in the average results, and the method of determining this is given in 6.1.2.

6.1.1 *Acceptability of results*

When only two results are obtained under repeatability conditions and their difference is less than or equal to r, the test operator may consider his work as being under control and may take the average of the two results as the estimated value of the property being tested.

If the two results differ by more than r, both shall be considered as suspect and at least three more results obtained. The difference between the most divergent result and the average of the remainder (including the first two) shall then be calculated and this difference compared with the repeatability of the method. If the difference is less than or equal to r, all the results shall be accepted.

This empirical rule is adopted for the sake of simplicity and corresponds to a slightly lower confidence level than the 95 % associated with two results (see 2.17).

If the difference exceeds the repeatability, the most divergent result shall be rejected and the procedure specified in this sub-clause repeated until an acceptable set of results is obtained. The average of the acceptable results shall be taken as the estimated value of the property. However, if two or more results from a total of not more than 20 have been rejected, the operating procedure and the apparatus shall be checked and a new series of tests made, if possible.

6.1.2 *Confidence limits*

When a single test operator, who is working within the precision limits of the method, obtains a series of n results under repeatability conditions, giving an average $\bar{x}_n$, it can be assumed with 95 % confidence that the true value μ of the characteristic lies within the following limits :

$$\mu = \bar{x}_n \pm \frac{1}{\sqrt{2}} \sqrt{R^2 - \left(1 - \frac{1}{n}\right) r^2} \text{ (double limit situation)} \quad \ldots (16)$$

Similarly, for the single limit situation, when only one upper or lower limit is fixed, it can be assumed with 95 % confidence that the true value μ of the characteristic is limited as follows :

$$\mu = \bar{x}_n + \frac{0{,}84}{\sqrt{2}} \sqrt{R^2 - \left(1 - \frac{1}{n}\right) r^2} \text{ (upper limit)} \quad \ldots (17)$$

or

$$\mu = \bar{x}_n - \frac{0{,}84}{\sqrt{2}} \sqrt{R^2 - \left(1 - \frac{1}{n}\right) r^2} \text{ (lower limit)} \quad \ldots (18)$$

NOTE – See annex H for an account of the statistical reasoning underlying these formulae.

However, since r for most test methods is much smaller than R, little improvement in the precision of the average is obtained by carrying out multiple testing under repeatability conditions.

NOTE – If the reproducibility R of a test method has been found considerably greater than twice the repeatability r, the reasons for the large value of the ratio R/r should be analysed and the method, if possible, should be improved.

6.2 Reproducibility R

6.2.1 *Acceptability of results*

The procedure specified in this sub-clause is intended for judging the acceptability, with respect to the reproducibility of the test method, of results obtained by different laboratories in normal day-to-day operations and transactions. In cases of dispute between a supplier and a recipient, the procedure specified in clause 7 shall be adopted.

When single results are obtained in two laboratories and their difference is less than or equal to R, the two results shall be considered as acceptable and their average, rather than either one separately, shall be considered as the estimated value of the tested property.

If the two results differ by more than R, both shall be considered as suspect. Each laboratory shall then obtain at least three other acceptable results (see 6.1.1).

In this case, the difference between the average of all acceptable results of each laboratory shall be judged for conformity using a new value R', instead of R, given by formula (19).

$$R' = \sqrt{R^2 - \left(1 - \frac{1}{2k_1} - \frac{1}{2k_2}\right) r^2} \quad \ldots (19)$$

where

R is the reproducibility of the method;

r is the repeatability of the method;

k_1 is the number of results of the first laboratory;

k_2 is the number of results of the second laboratory.

If circumstances arise in which more than two laboratories supply single results, the difference between the most divergent result and the average of the remainder shall be compared with the reproducibility of the method. If this difference is equal to, or less than, the reproducibility, all the results shall be regarded as acceptable and their average taken as the estimated value of the property.

If, however, the difference is greater than the reproducibility, the most divergent result shall be rejected and the procedure specified in this sub-clause repeated until an acceptable set of results is obtained. Then the average of these results shall be taken as the estimated value of the property. However, if two or more results from a total of not more than 20 have been rejected, the operating procedure and the apparatus shall be checked and a new series of tests made if possible.

6.2.2 *Confidence limits*

When k laboratories obtain single results under reproducibility conditions, giving an average $\overline{X}_k$, it may be assumed with 95 % confidence that the true value μ of the characteristic lies within the following limits :

$$\mu = \overline{X}_k \pm \frac{R}{\sqrt{2k}} \text{ (double limit situation)} \quad \ldots (20)$$

Similarly for the single limit situation, when only one upper or lower limit is fixed, it may be assumed with 95 % confidence that the true value μ of the characteristic is limited as follows :

$$\mu = \overline{X}_k + \frac{0{,}84\ R}{\sqrt{2k}} \text{ (upper limit)} \quad \ldots (21)$$

and

$$\mu = \overline{X}_k - \frac{0{,}84\ R}{\sqrt{2k}} \text{ (lower limit)} \quad \ldots (22)$$

NOTE — See annex H for an account of the statistical reasoning underlying these formulae.

These formulae also allow a given laboratory ($k = 1$) to determine the confidence level that can be assigned to a single result by comparison with the true value μ.

7 SPECIFICATIONS

7.1 Aim of specifications

The purpose of a specification should be to fix a limit or limits to the true value of the property considered. In practice, however, this true value can never be established exactly. The property is measured in the laboratory by applying a standard test method, the results of which may show some scattering as defined by the repeatability and reproducibility limits. There is therefore some uncertainty as to the true value of the tested property.

7.2 Construction of specifications

Specifications deal usually with limits for the values of the properties. To avoid uncertainty, such limits should normally be expressed as "not less than" or "not greater than". Limits are of two types :

— a double limit, upper and lower, for example viscosity not less than 5 mm^2/s [1] and not greater than 10 mm^2/s; boiling point 100 ± 0,5 °C;

— a single limit, upper or lower, for example sulphur content not greater than 2 %; lead content not greater than 3,0 g/l; solubility of bitumen not less than 99 %.

The single limit situation becomes relevant when, as in most cases, there is an additional implied limit which effectively converts it into a double limit situation. This is illustrated by the examples above in which the additional implied limits are 0 %, 0 g/l, and 100 % respectively. In cases of a true single limit situation, for example flash point not less than 60 °C, the following considerations do not apply.

The value chosen for a specification limit shall take into account the reproducibility of the test method adopted, as follows :

For a double limit (A_1 and A_2), the specified range, (stated or implied) shall be not less than four times the reproducibility R, i.e. $A_2 - A_1 \geqslant 4\ R$.

For a single limit (A_1 or A_2), the specified limit shall be not less than twice the reproducibility R, i.e. $A_1 \geqslant 2\ R$ or $A_2 \geqslant 2\ R$.

The requirements of this International Standard apply to specifications drawn up in accordance with these principles.

NOTE — In cases where, for practical reasons, the value of $A_2 - A_1$ is less than 4 R, the results obtained will be of doubtful significance in determining whether a sample does or does not satisfy the requirements of the specification [2]. In such cases, one or both of the following courses should be adopted :

a) The specification limits should be examined to see whether they can be widened to fit in with the precision of the test method.

b) The test method should be examined to see whether the precision can be improved, or an alternative method adopted with an improved precision, to fit in with the desired specification limits.

8 QUALITY CONTROL AGAINST SPECIFICATIONS

8.0 Introduction

This clause provides general information to allow the supplier or the recipient to judge the quality of a product with regard to the specification when a single result is available. If it is necessary for the recipient to take action after examining this result, the procedure specified in clause 9 shall be adopted.

8.1 Testing margin at the supplier

A supplier who has no other source of information on the true value of a characteristic than a single result shall consider that the product meets the specification limit, with 95 % confidence, only if the result X is such that

in the case of a single upper limit A_1,

$$X \leqslant A_1 - \frac{0{,}84\ R}{\sqrt{2}} \quad \ldots (23)$$

1) 1 mm^2/s = 1 cSt.

2) According to statistical reasoning, it is desirable for $A_2 - A_1$ to be considerably greater than 4R.

in the case of a single lower limit A_2,

$$X \geqslant A_2 + \frac{0{,}84\,R}{\sqrt{2}} \qquad \ldots (24)$$

in the case of a double limit (A_1 and A_2), both these conditions are satisfied.

(See 6.1.2 and 6.2.2.)

NOTE — The use of equations 23 and 24 is for the guidance of the supplier and is not to be interpreted as an obligation. A reported value between the specification value and the limit from equation 23/24 is not proof of non-compliance.

8.2 Testing margin at the recipient

A recipient who has no other source of information on the true value of a characteristic than a single result shall consider that the product fails the specification limit, with 95 % confidence, only if the result X is such that

in the case of a single upper limit A_1,

$$X > A_1 + \frac{0{,}84\,R}{\sqrt{2}} \qquad \ldots (25)$$

in the case of a single lower limit A_2,

$$X < A_2 - \frac{0{,}84\,R}{\sqrt{2}} \qquad \ldots (26)$$

in the case of a double limit (A_1 and A_2), either of these conditions applies.

9 ACCEPTANCE AND REJECTION RULES IN CASE OF DISPUTE

9.0 If it is not possible for the supplier and the recipient to reach agreement about the quality of the product, on the basis of their existing results, then the following procedures shall be adopted.

9.1 Each laboratory shall reject its original results and obtain at least three other acceptable results on the check sample to ensure that the work has been carried out under repeatability conditions. The average of the acceptable results in each laboratory shall then be computed, divergent results being discarded as indicated in 6.1.1.

Let

$\overline{X}_S$ be the average of supplier;

$\overline{X}_R$ be the average of recipient;

A_1 be the upper limit of the specification;

A_2 be the lower limit of the specification;

where

$\overline{X}_S \leqslant A_1$ or $\geqslant A_2$;

$\overline{X}_R > A_1$ or $< A_2$.

This means that $\overline{X}_S$ and $\overline{X}_R$ should be compared as follows with A_1 and A_2.

9.1.1 If

$$\frac{\overline{X}_S + \overline{X}_R}{2} \leqslant A_1 \text{ or } \geqslant A_2,$$

product accepted if $|\overline{X}_S - \overline{X}_R| \leqslant 0{,}84\,R'$ (for R', see 6.2.1), possible dispute if $|\overline{X}_S - \overline{X}_R| > 0{,}84\,R'$.

NOTE — In this case it cannot be stated with confidence whether the product does or does not comply with the specification limit; hence, resolution of the dispute may be by negotiation.

9.1.2 If

$$\frac{\overline{X}_S + \overline{X}_R}{2} > A_1 \text{ or } < A_2,$$

dispute whatever the difference $\overline{X}_S - \overline{X}_R$.

9.2 In case of dispute, the two laboratories shall contact each other and compare their operating procedures and apparatus. Following these investigations, a correlation test between the two laboratories shall be carried out on the two check samples. The average of at least three acceptable results shall be computed, in each laboratory, and these averages compared as indicated in 9.1 and above.

9.3 If the disagreement remains, a third laboratory (neutral, expert and accepted by the two parties) shall be invited to carry out the test using a third sample. Suppose $\overline{X}_E$ is the average of the three acceptable results of the third laboratory. If the difference between the most divergent laboratory average and the average of the two other laboratory averages is less than or equal to R, the following procedure shall be adopted :

9.3.1 If

$$\frac{\overline{X}_S + \overline{X}_R + \overline{X}_E}{3} \leqslant A_1 \text{ or } \geqslant A_2,$$

product accepted.

9.3.2 If

$$\frac{\overline{X}_S + \overline{X}_R + \overline{X}_E}{3} > A_1 \text{ or } < A_2,$$

product rejected.

9.4 If the difference between the most divergent laboratory average and the average, $\overline{X}$, of the two other laboratory averages is more than R, the following procedure shall be adopted :

9.4.1 If

$$\overline{X} \leqslant A_1 \text{ or } \geqslant A_2,$$

product accepted.

9.4.2 If

$$\overline{X} > A_1 \text{ or } < A_2,$$

product rejected.

ANNEX A

DETERMINATION OF NUMBER OF SAMPLES REQUIRED (see 3.3)

TABLE 11

L = number of participating laboratories

$P = \frac{\text{interaction variance component}}{\text{repeats variance component}}$

$Q = \frac{\text{laboratories variance component}}{\text{repeats variance component}}$

$L = 5$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	4									
1	5									
2	6	11								
3	6	9								
4	7	8	16							
5	7	8	12							
6	7	8	11	19						
7	7	8	10	15						
8	7	8	9	13						
9	7	8	9	11	17					

$L = 6$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	3									
1	4	11								
2	5	7								
3	5	7	14							
4	5	6	10							
5	6	6	8	15						
6	6	6	8	11						
7	6	6	7	10	15					
8	6	6	7	9	12					
9	6	6	7	8	10	15				

$L = 7$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	3									
1	4	7								
2	4	6	17							
3	4	5	9							
4	5	5	7	13						
5	5	5	6	9	19					
6	5	5	6	8	12					
7	5	5	6	7	10	15				
8	5	5	6	7	8	12	20			
9	5	5	6	6	8	10	14			

$L = 8$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	3									
1	3	5								
2	4	5	9							
3	4	5	7	14						
4	4	4	6	9	20					
5	4	4	5	7	11					
6	4	4	5	6	8	13				
7	4	4	5	6	7	10	16			
8	4	5	5	6	6	8	11	18		
9	4	5	5	5	6	7	9	13		

$L = 9$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2									
1	3	4								
2	3	4	7							
3	3	4	5	9						
4	4	4	5	6	11					
5	4	4	5	6	7	12				
6	4	4	4	5	6	9	14			
7	4	4	4	5	6	7	10	15		
8	4	4	4	5	5	6	8	10	16	
9	4	4	4	5	5	6	7	8	11	18

$L = 10$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	8								
1	3	4	11							
2	3	4	5	12						
3	3	3	4	6	13					
4	3	4	4	5	7	14				
5	3	4	4	5	6	8	14			
6	3	4	4	4	5	6	9	14		
7	3	4	4	4	5	6	7	9	14	
8	3	4	4	4	5	5	6	7	10	14
9	4	4	4	4	4	5	6	6	8	10

$L = 11$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	4								
1	2	3	5							
2	3	3	3	7						
3	3	3	4	5	8					
4	3	3	4	4	6	8	18			
5	3	3	4	4	5	6	9	15		
6	3	3	3	4	4	5	6	9	14	
7	3	3	3	4	4	5	5	7	9	13
8	3	3	3	4	4	4	5	6	7	9
9	3	3	3	4	4	4	5	5	6	7

$L = 12$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	4								
1	2	3	5							
2	2	3	4	6	14					
3	3	3	3	4	6	11				
4	3	3	3	4	5	6	9			
5	3	3	3	4	4	5	6	9	16	
6	3	3	3	3	4	4	5	6	9	13
7	3	3	3	3	4	4	5	5	6	8
8	3	3	3	3	4	4	4	5	5	6
9	3	3	3	3	3	4	4	4	5	6

$L = 13$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	3								
1	2	3	4	12						
2	2	3	3	4	8					
3	2	3	3	4	5	7	14			
4	3	3	3	3	4	5	7	10		
5	3	3	3	3	4	4	5	6	9	15
6	3	3	3	3	3	4	4	5	6	8
7	3	3	3	3	3	4	4	4	5	6
8	3	3	3	3	3	3	4	4	5	5
9	3	3	3	3	3	3	4	4	4	5

$L = 14$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	3								
1	2	2	3	7						
2	2	2	3	4	6	12				
3	2	2	3	3	4	5	8	18		
4	2	3	3	3	3	4	5	7	11	
5	2	3	3	3	3	4	4	5	6	8
6	3	3	3	3	3	3	4	4	5	6
7	3	3	3	3	3	3	3	4	4	5
8	3	3	3	3	3	3	3	4	4	4
9	3	3	3	3	3	3	3	3	4	4

$L = 15$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	2	2	13							
1	2	2	3	5	19					
2	2	2	3	3	4	7				
3	2	2	3	3	3	4	6	9		
4	2	2	3	3	3	4	4	5	7	10
5	2	2	3	3	3	3	4	4	5	6
6	2	2	3	3	3	3	3	4	4	5
7	2	2	3	3	3	3	3	3	4	4
8	2	2	3	3	3	3	3	3	3	4
9	2	2	3	3	3	3	3	3	3	3

$L = 16$

P \ Q	0	1	2	3	4	5	6	7	8	9
0	1	2	5							
1	2	2	3	4	8					
2	2	2	2	3	4	5	9			
3	2	2	2	3	3	4	4	6	9	
4	2	2	2	3	3	3	4	4	5	6
5	2	2	2	3	3	3	3	4	4	5
6	2	2	2	2	3	3	3	3	4	4
7	2	2	2	2	3	3	3	3	3	4
8	2	2	2	2	3	3	3	3	3	3
9	2	2	2	2	3	3	3	3	3	3

ANNEX B

DERIVATION OF FORMULA FOR CALCULATING THE NUMBER OF SAMPLES REQUIRED (see 3.3)

An analysis of variance is carried out on the results of the pilot programme. This yields rough estimates of the three components of variance, namely :

σ_0^2 for repeats,

σ_1^2 for laboratories × samples interaction,

σ_2^2 for laboratories.

Substituting the above in formula (14) (see 5.4.3.2) for calculating the reproducibility degrees of freedom, this becomes

$$\frac{(1+p+q)^2}{\nu} = \frac{\left[\left(\frac{1}{2}+p\right)\Big/ S+q\right]^2}{L-1} + \frac{(S-1)\left(\frac{1}{2}+p\right)^2}{S^2\,(L-1)} + \frac{1}{4LS} \qquad \ldots (27)$$

where

p is the ratio σ_1^2/σ_0^2;

q is the ratio σ_2^2/σ_0^2;

ν is the reproducibility degrees of freedom;

L is the number of laboratories;

S is the number of samples.

The formula rearranges into the form

$$aS + b = 0$$

where

$$a = \nu q^2 - (1+p+q)^2\,(L-1)$$

$$b = \nu\left[\left(2q+\frac{1}{2}+p\right)\left(\frac{1}{2}+p\right) + 0{,}25\,(L-1)/L\right]$$

Therefore

$$S = -\frac{b}{a} \qquad \ldots (28)$$

gives the values of S for given values of L, p, q and ν.

ANNEX C

NOTATION AND TESTS

Throughout this International Standard the following notation is used :

S is the number of samples;

L is the number of laboratories;

i is the suffix denoting laboratory number;

j is the suffix denoting sample number;

x is an individual test result;

a is the sum of duplicate test results;

e is the difference between duplicate test results.

C.1 ARRAY OF DUPLICATE RESULTS FROM EACH OF L LABORATORIES ON S SAMPLES AND CORRESPONDING MEANS m_j

TABLE 12

Laboratory	Sample 1	2	j	S
1	x_{111} x_{112}	x_{121} x_{122}	x_{1j1} x_{1j2}	x_{1S1} x_{1S2}
2	x_{211} x_{212}	x_{221} x_{222}	x_{2j1} x_{2j2}	x_{2S1} x_{2S2}
i	x_{i11} x_{i12}	x_{i21} x_{i22}	x_{ij1} x_{ij2}	x_{iS1} x_{iS2}
L	x_{L11} x_{L12}	x_{L21} x_{L22}	x_{Lj1} x_{Lj2}	x_{LS1} x_{LS2}
Total	g_1	g_2	g_j	g_S
Mean	m_1	m_2	m_j	m_S

NOTE — If a transformation $y = F(x)$ of the reported data is necessary (see 4.1), then corresponding symbols y_{ij1} and y_{ij2} are used in place of x_{ij1} and x_{ij2}.

C.2 ARRAY OF SUMS OF DUPLICATE RESULTS OF LABORATORY TOTALS h_j AND SAMPLE TOTALS g_j

TABLE 13

Laboratory	Sample 1	2	j	S	Total
1	a_{11}	a_{12}	a_{1j}	a_{1S}	h_1
2	a_{21}	a_{22}	a_{2j}	a_{2S}	h_2
i	a_{i1}	a_{i2}	a_{ij}	a_{iS}	h_i
L	a_{L1}	a_{L2}	a_{Lj}	a_{LS}	h_L
Total	g_1	g_2	g_j	g_S	T

$a_{ij} = x_{ij1} + x_{ij2}$ (or $a_{ij} = y_{ij1} + y_{ij2}$, if a transformation has been used)

$e_{ij} = x_{ij1} - x_{ij2}$

$$g_j = \sum_{i=1}^{L} a_{ij}$$

$m_j = g_j/2L$

$$h_i = \sum_{j=1}^{S} a_{ij}$$

$$T = \sum_{i=1}^{L} h_i = \sum_{j=1}^{S} g_j$$

NOTE – If any results are missing from the complete array, then the divisors in the expression for m_j will be correspondingly reduced.

C.3 SUMS OF SQUARES AND VARIANCES (see 4.1)

Repeats variance for sample j

$$d_j^2 = \frac{\sum_{i=1}^{L} e_{ij}^2}{2L} \qquad \ldots (29)$$

NOTE – If either or both of a laboratory/sample pair of results is missing, the corresponding term in the numerator is omitted, and the factor $2L$ is reduced by 2.

Within sample variance for sample j

$$W_j^2 = \left[\sum_{j=1}^{S} (x_{ij1}^2 + x_{ij2}^2) - \frac{g_j^2}{2L}\right] \Big/ (2L - 1)$$

NOTE – The factors $2L$ and $(2L - 1)$ are reduced by 1 for every missing value in the sample.

Laboratories variance for sample j

$$D_j^2 = \frac{1}{K_j}\left[W_j^2 + (K_j - 1)\, d_j^2\right] \qquad \ldots (30)$$

where

$$K_j = \left(S_j^2 - \sum_{i=1}^{L} n_{ij}^2\right) \Big/ [S_j(S_j - 1)]$$

n_{ij} is the number of results obtained by laboratory i from sample j;

S_j is the total number of results obtained from sample j.

C.4 COCHRAN'S TEST

$$\text{Cochran's criterion} = \frac{\text{largest } e_{ij}^2}{\sum_{i=1}^{L} \sum_{j=1}^{S} e_{ij}^2} \qquad \ldots (31)$$

The associated number K is the number of complete pairs in the array of C.1. The critical 1 % values are set out in table 17 (annex D).

C.5 DIXON'S TESTS

Table 14 shows the corresponding r ratio for any sample (or for the laboratory totals) the pair sums a_{ij} of which are arranged in ascending order, $a_1, a_2, a_3, \ldots, a_{L-1}, a_L$.

TABLE 14

Ratio	For low values	For high values
r_{10}	$\frac{a_2 - a_1}{a_L - a_1}$	$\frac{a_L - a_{L-1}}{a_L - a_1}$
r_{11}	$\frac{a_2 - a_1}{a_{L-1} - a_1}$	$\frac{a_L - a_{L-1}}{a_L - a_2}$
r_{12}	$\frac{a_2 - a_1}{a_{L-2} - a_1}$	$\frac{a_L - a_{L-1}}{a_L - a_3}$
r_{20}	$\frac{a_3 - a_1}{a_L - a_1}$	$\frac{a_L - a_{L-2}}{a_L - a_1}$
r_{21}	$\frac{a_3 - a_1}{a_{L-1} - a_1}$	$\frac{a_L - a_{L-2}}{a_L - a_2}$
r_{22}	$\frac{a_3 - a_1}{a_{L-2} - a_1}$	$\frac{a_L - a_{L-2}}{a_L - a_3}$

The critical 1 % values for these ratios are given in table 18 (annex D).

ANNEX D

EXAMPLE OF PROCEDURE FOR INSPECTION OF RESULTS OF TEST FOR DETERMINATION OF BROMINE NUMBER

D.1 BROMINE NUMBER FOR LOW BOILING SAMPLES

TABLE 15

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	1,9	64,5	0,80	3,7	11,0	46,1	114,8	1,2
	2,1	65,5	0,78	3,8	11,1	46,5	114,2	1,2
B	1,7	65,4	0,69	3,7	11,1	50,3	114,5	1,2
	1,8	66,0	0,72	3,7	11,0	49,9	114,3	1,2
C	1,8	63,5	0,76	3,5	10,4	48,5	112,4	1,3
	1,8	63,8	0,76	3,5	10,5	48,2	112,7	1,3
D	4,1	63,6	0,80	4,0	10,8	49,6	108,8	1,0
	4,0	63,9	0,80	3,9	10,8	49,9	108,2	1,1
E	2,1	63,9	0,83	3,7	10,9	47,4	115,6	1,3
	1,8	63,7	0,83	3,7	11,1	47,6	115,1	1,4
F	1,8	70,7	0,72	3,4	11,5	49,1	121,0	1,4
	1,7	69,7	0,64	3,6	11,2	47,9	117,9	1,4
G	1,9	63,8	0,77	3,5	10,6	46,1	114,1	1,1
	2,2	63,6	0,59	3,5	10,6	45,5	112,8	0,93
H	2,0	66,5	0,78	3,2	10,7	49,6	114,8	1,1
	1,8	65,5	0,71	3,5	10,7	48,5	114,5	1,0
J	2,1	68,2	0,81	4,0	11,1	49,1	115,7	1,4
	2,1	65,3	0,81	3,7	11,1	47,9	113,9	1,4

D.2 CUBE ROOT OF BROMINE NUMBER FOR LOW BOILING SAMPLES

TABLE 16

Laboratory	Sample							
	1	2	3	4	5	6	7	8
A	1,239	4,010	0,928	1,547	2,224	3,586	4,860	1,063
	1,281	4,031	0,921	1,560	2,231	3,596	4,852	1,063
B	1,193	4,029	0,884	1,547	2,231	3,691	4,860	1,063
	1,216	4,041	0,896	1,547	2,224	3,682	4,853	1,063
C	1,216	3,990	0,913	1,518	2,183	3,647	4,826	1,091
	1,216	3,996	0,913	1,518	2,190	3,639	4,830	1,091
D	1,601	3,992	0,928	1,587	2,210	3,674	4,774	1,000
	1,587	3,998	0,928	1,574	2,210	3,682	4,765	1,032
E	1,281	3,998	0,940	1,547	2,217	3,619	4,871	1,091
	1,216	3,994	0,940	1,547	2,231	3,624	4,864	1,119
F	1,216	4,135	0,896	1,504	2,257	3,662	4,946	1,119
	1,193	4,115	0,862	1,533	2,237	3,632	4,903	1,119
G	1,239	3,996	0,917	1,518	2,197	3,586	4,850	1,032
	1,301	3,992	0,839	1,518	2,197	3,570	4,832	0,976
H	1,260	4,051	0,921	1,474	2,204	3,674	4,860	1,032
	1,216	4,031	0,892	1,518	2,204	3,647	4,856	1,000
J	1,281	4,086	0,932	1,587	2,231	3,662	4,873	1,119
	1,281	4,027	0,932	1,547	2,231	3,632	4,847	1,119

D.3 UPPER 1 % LEVELS OF COCHRAN'S CRITERION FOR VALUES OF K FROM 12 TO 120

TABLE 17

K	Cochran's criterion	K	Cochran's criterion	K	Cochran's criterion
12	0,652 8	45	0,269 3	85	0,163 6
15	0,574 7	50	0,248 4	90	0,156 1
20	0,479 9	55	0,230 6	95	0,149 3
24	0,424 7	60	0,215 1	100	0,143 0
25	0,413 0	65	0,202 4	105	0,137 3
30	0,363 2	70	0,191 0	110	0,132 0
35	0,325 0	75	0,180 9	115	0,127 1
40	0,294 0	80	0,171 8	120	0,122 5

If the calculated ratio for K ranges is less than the nearest critical ratio with more ranges, or if it is greater than the nearest critical ratio with less ranges, no calculation will be required. In the latter case a rejection will be made and in the former, none.

If intermediate values are required, these may be obtained by linear interpolation of the reciprocals of the results underlined in table 17, which are exact values. Values for over 120 ranges may be found by extrapolating from the reciprocals of the results for 60 and 120.

D.4 CRITICAL VALUES OF DIXON'S 1 %* TESTS FOR VALUES OF n FROM 3 TO 30

TABLE 18

n	r_{10}	r_{11}	r_{21}	r_{22}
3	0,994			
4	0,926			
5	0,821			
6	0,740			
7	0,680			
8		0,725		
9		0,677		
10		0,639		
11			0,713	
12			0,675	
13			0,649	
14				0,674
15				0,647
16				0,624
17				0,605
18				0,589
19				0,575
20				0,562
21				0,551
22				0,541
23				0,532
24				0,524
25				0,516
26				0,508
27				0,501
28				0,495
29				0,489
30				0,483

* The tabulated values refer to the 0,5 % probability level, but since the test is carried out for both ends of the range, it becomes a 1 % rejection test.

D.5 CRITICAL VALUES OF *t*

TABLE 19

Degrees of freedom	Per Cent Probability Level						
	50	40	30	20	10	5	1
1	1,000	1,376	1,963	3,078	6,314	12,706	63,657
2	0,816	1,061	1,386	1,886	2,920	4,303	9,925
3	0,765	0,978	1,250	1,638	2,353	3,182	5,841
4	0,741	0,941	1,190	1,533	2,132	2,776	4,604
5	0,727	0,920	1,156	1,476	2,015	2,571	4,032
6	0,718	0,906	1,134	1,440	1,943	2,447	3,707
7	0,711	0,896	1,119	1,415	1,895	2,365	3,499
8	0,706	0,889	1,108	1,397	1,860	2,306	3,355
9	0,703	0,883	1,100	1,383	1,833	2,262	3,250
10	0,700	0,879	1,093	1,372	1,812	2,228	3,169
11	0,697	0,876	1,088	1,363	1,796	2,201	3,106
12	0,695	0,873	1,083	1,356	1,782	2,179	3,055
13	0,694	0,870	1,079	1,350	1,771	2,160	3,012
14	0,692	0,868	1,076	1,345	1,761	2,145	2,977
15	0,691	0,866	1,074	1,341	1,753	2,131	2,947
16	0,690	0,865	1,071	1,337	1,746	2,120	2,921
17	0,689	0,863	1,069	1,333	1,740	2,110	2,898
18	0,688	0,862	1,067	1,330	1,734	2,101	2,878
19	0,688	0,861	1,066	1,328	1,729	2,093	2,861
20	0,687	0,860	1,064	1,325	1,725	2,086	2,845
21	0,686	0,859	1,063	1,323	1,721	2,080	2,831
22	0,686	0,858	1,061	1,321	1,717	2,074	2,819
23	0,685	0,858	1,060	1,319	1,714	2,069	2,807
24	0,685	0,857	1,059	1,318	1,711	2,064	2,797
25	0,684	0,856	1,058	1,316	1,708	2,060	2,787
26	0,684	0,856	1,058	1,315	1,706	2,056	2,779
27	0,684	0,855	1,057	1,314	1,703	2,052	2,771
28	0,683	0,855	1,056	1,313	1,701	2,048	2,763
29	0,683	0,854	1,055	1,311	1,699	2,045	2,756
30	0,683	0,854	1,055	1,310	1,697	2,042	2,750
40	0,681	0,851	1,050	1,303	1,684	2,021	2,704
50	0,680	0,849	1,048	1,299	1,676	2,008	2,678
60	0,679	0,848	1,046	1,296	1,671	2,000	2,660
120	0,677	0,845	1,041	1,289	1,658	1,980	2,617
∞	0,674	0,842	1,036	1,282	1,645	1,960	2,576