Designation: D6300 – 24

An American National Standard

Standard Practice for Determination of Precision and Bias Data for Use in Test Methods for Petroleum Products, Liquid Fuels, and Lubricants[1]

This standard is issued under the fixed designation D6300; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

INTRODUCTION

Both Research Report RR:D02-1007,[2] *Manual on Determining Precision Data for ASTM Methods on Petroleum Products and Lubricants*[2] and the ISO 4259, benefitted greatly from more than 50 years of collaboration between ASTM and the Institute of Petroleum (IP) in the UK. The more recent work was documented by the IP and has become ISO 4259.

ISO 4259 encompasses both the determination of precision and the application of such precision data. In effect, it combines the type of information in RR:D02-1007[2] regarding the determination of the precision estimates and the type of information in Practice D3244 for the utilization of test data. The following practice, intended to replace RR:D02-1007,[2] differs slightly from related portions of the ISO standard.

1. Scope*

1.1 This practice covers the necessary preparations and planning for the conduct of interlaboratory programs for the development of estimates of precision (determinability, repeatability, and reproducibility) and of bias (absolute and relative), and further presents the standard phraseology for incorporating such information into standard test methods.

1.2 This practice is generally limited to homogeneous petroleum products, liquid fuels, and lubricants with which serious sampling problems (such as heterogeneity or instability) do not normally arise.

1.3 This practice may not be suitable for products with sampling problems as described in 1.2, solid or semisolid products such as petroleum coke, industrial pitches, paraffin waxes, greases, or solid lubricants when the heterogeneous properties of the substances create sampling problems. In such instances, consult a trained statistician.

1.4 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[3]

D3244 Practice for Utilization of Test Data to Determine Conformance with Specifications

D3606 Test Method for Determination of Benzene and Toluene in Spark Ignition Fuels by Gas Chromatography

D6708 Practice for Statistical Assessment and Improvement of Expected Agreement Between Two Test Methods that Purport to Measure the Same Property of a Material

D7915 Practice for Application of Generalized Extreme Studentized Deviate (GESD) Technique to Simultaneously Identify Multiple Outliers in a Data Set

E29 Practice for Using Significant Digits in Test Data to Determine Conformance with Specifications

E177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods

E456 Terminology Relating to Quality and Statistics

E691 Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method

[1] This practice is under the jurisdiction of ASTM Committee D02 on Petroleum Products, Liquid Fuels, and Lubricantsand is the direct responsibility of Subcommittee D02.94 on Coordinating Subcommittee on Quality Assurance and Statistics.

Current edition approved March 1, 2024. Published March 2024. Originally approved in 1998. Last previous edition approved in 2023 as D6300 – 23a. DOI: 10.1520/D6300-24.

[2] Supporting data have been filed at ASTM International Headquarters and may be obtained by requesting Research Report RR:D02-1007. Contact ASTM Customer Service at service@astm.org.

[3] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

*A Summary of Changes section appears at the end of this standard

2.2 *ISO Standards:*

ISO 4259 Petroleum Products-Determination and Application of Precision Data in Relation to Methods of Test[4]

3. Terminology

3.1 *Definitions:*

3.1.1 *analysis of variance (ANOVA), n*—technique that enables the total variance of a method to be broken down into its component factors. **ISO 4259**

3.1.2 *bias, n*—the difference between the expectation of the test results and an accepted reference value.

3.1.2.1 *Discussion*—The term "expectation" is used in the context of statistics terminology, which implies it is a "statistical expectation." **E177**

3.1.3 *between-method bias (relative bias), n*—a quantitative expression for the mathematical correction that can statistically improve the degree of agreement between the expected values of two test methods which purport to measure the same property. **D6708**

3.1.4 *degrees of freedom, n*—the divisor used in the calculation of variance, one less than the number of independent results.

3.1.4.1 *Discussion*—This definition applies strictly only in the simplest cases. Complete definitions are beyond the scope of this practice. **ISO 4259**

3.1.5 *determinability, n*—a quantitative measure of the variability associated with the same operator in a given laboratory obtaining successive determined values using the same apparatus for a series of operations leading to a single result; it is defined as the difference between two such single determined values that would be exceeded about 5 % of the time (one case in 20 in the long run) in the normal and correct operation of the test method.

3.1.5.1 *Discussion*—This definition implies that two determined values, obtained under determinability conditions, which differ by more than the determinability value should be considered suspect. If an operator obtains more than two determinations, then it would usually be satisfactory to check the most discordant determination against the mean of the remainder, using determinability as the critical difference (**1**).[5]

3.1.6 *mean square, n—in analysis of variance*, sum of squares divided by the degrees of freedom. **ISO 4259**

3.1.7 *normal distribution, n*—the distribution that has the probability function x, such that, if x is any real number, the probability density is

$$f(x) = (1/\sigma)(2\pi)^{-1/2}\exp[-(x-\mu)^2/2\sigma^2] \quad (1)$$

NOTE 1—μ is the true value and σ is the standard deviation of the normal distribution (σ > 0). **ISO 4259**

3.1.8 *outlier, n*—a result far enough in magnitude from other results to be considered not a part of the set. **RR:D02–1007[2]**

3.1.9 *precision, n*—the degree of agreement between two or more results on the same property of identical test material. In this practice, precision statements are framed in terms of *repeatability* and *reproducibility* of the test method.

3.1.9.1 *Discussion*—The testing conditions represented by repeatability and reproducibility should reflect the normal extremes of variability under which the test is commonly used. Repeatability conditions are those showing the least variation; reproducibility, the usual maximum degree of variability. Refer to the definitions of each of these terms for greater detail. **RR:D02–1007[2]**

3.1.10 *random error, n*—the chance variation encountered in all test work despite the closest control of variables. **RR:D02–1007[2]**

3.1.11 *repeatability (a.k.a. Repeatability Limit), n*—a quantitative expression for the random error associated with the difference between two independent results obtained under repeatability conditions that would be exceeded about 5 % of the time (one case in 20 in the long run) in the normal and correct operation of the test method.

3.1.11.1 *Discussion*—Interpret as the limit value the absolute difference between two single test results obtained under repeatability conditions is expected to exceed with an approximate probability of 5 %.

3.1.11.2 *Discussion*—The difference is related to the repeatability standard deviation but it is not the standard deviation or its estimate.

3.1.11.3 *Discussion*—In 3.1.11 and 3.1.13, the term "probability" quantifies the likelihood of repeatability or reproducibility limit exceedance for the difference between a single pair of results obtained under the respective conditions. The "one case in 20 in the long run" in the parenthesis is not to be interpreted as one case in every 20, but it is over the long run. The long run concept can be illustrated using 10 cases out of 200, or 100 cases out of 2000, or 1000 cases in 20 000. The lowest numerical values of one case in 20 is used here.

3.1.11.4 *Discussion*—The "one case in 20" is a legacy term that was carried over from RR:D02-1007 in the original development of Practice D6300. **RR:D02–1007[2]**

3.1.12 *repeatability conditions, n*—conditions where independent test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time. **E177**

3.1.13 *reproducibility (a.k.a. Reproducibility Limit), n*—a quantitative expression for the random error associated with the difference between two independent results obtained under reproducibility conditions that would be exceeded about 5 % of the time (one case in 20 in the long run) in the normal and correct operation of the test method.

3.1.13.1 *Discussion*—Interpret as the limit value the absolute difference between two single test results obtained under reproducibility conditions is expected to exceed with an approximate a probability of 5 %.

3.1.13.2 *Discussion*—The difference is related to the reproducibility standard deviation but is not the standard deviation or its estimate. **RR:D02–1007[2]**

3.1.13.3 *Discussion*—In those cases where the normal use of the test method does not involve sending a sample to a testing laboratory, either because it is an in-line test method or

[4] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

[5] The bold numbers in parentheses refers to the list of references at the end of this standard.

because of serious sample instabilities or similar reasons, the precision test for obtaining reproducibility may allow for the use of apparatus from the participating laboratories at a common site (several common sites, if feasible). The statistical analysis is not affected thereby. However, the interpretation of the reproducibility value will be affected since the test data is collected under intermediate precision conditions as defined in Practice E177, and therefore, the precision statement shall, in this case, state the conditions to which the reproducibility value applies, and label this precision in a manner consistent with how the test data is obtained.

NOTE 2—The reproducibility precision outcome from 3.1.13.3 is a form of Intermediate Precision as defined in Practice E177.

3.1.14 *reproducibility conditions, n*—conditions where independent test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.

NOTE 3—Different laboratory by necessity means a different operator, different equipment, and different location and under different supervisory control. **E177**

3.1.15 *standard deviation, n*—measure of the dispersion of a series of results around their mean, equal to the square root of the variance and estimated by the positive square root of the mean square. **ISO 4259**

3.1.16 *sum of squares, n—in analysis of variance*, sum of squares of the differences between a series of results and their mean. **ISO 4259**

3.1.17 *variance, n*—a measure of the dispersion of a series of accepted results about their average. It is equal to the sum of the squares of the deviation of each result from the average, divided by the number of degrees of freedom. **RR:D02–1007[2]**

3.1.18 *variance, between-laboratory, n*—that component of the overall variance due to the difference in the mean values obtained by different laboratories. **ISO 4259**

3.1.18.1 *Discussion*—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. Differences in operator technique, instrumentation, environment, and sample "as received" are among the factors that can affect the between laboratory variance. There is a corresponding definition for between-operator variance.

3.1.18.2 *Discussion*—The term "between-laboratory" is often shortened to "laboratory" when used to qualify representative parameters of the dispersion of the population of results, for example as "laboratory variance."

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *determination, n*—the process of carrying out a series of operations specified in the test method whereby a single value is obtained.

3.2.2 *operator, n*—a person who carries out a particular test.

3.2.3 *probability density function, n*—function which yields the probability that the random variable takes on any one of its admissible values; here, we are interested only in the normal probability.

3.2.4 *result, n*—the final value obtained by following the complete set of instructions in the test method.

3.2.4.1 *Discussion*—It may be obtained from a single determination or from several determinations, depending on the instructions in the method. When rounding off results, the procedures described in Practice E29 shall be used.

4. Summary of Practice

4.1 A draft of the test method is prepared and a pilot program can be conducted to verify details of the procedure and to estimate roughly the precision of the test method.

4.1.1 If the responsible committee decides that an interlaboratory study for the test method is to take place at a later point in time, an interim repeatability is estimated by following the requirements in 6.2.1.

4.2 A plan is developed for the interlaboratory study using the number of participating laboratories to determine the number of samples needed to provide the necessary degrees of freedom. Samples are acquired and distributed. The interlaboratory study is then conducted on an agreed draft of the test method.

4.3 The data are summarized and analyzed. Any dependence of precision on the level of test result is removed by transformation. The resulting data are inspected for uniformity and for outliers. Any missing and rejected data are estimated. The transformation is confirmed. Finally, an analysis of variance is performed, followed by calculation of repeatability, reproducibility, and bias. When it forms a necessary part of the test procedure, the determinability is also calculated.

5. Significance and Use

5.1 ASTM test methods are frequently intended for use in the manufacture, selling, and buying of materials in accordance with specifications and therefore should provide such precision that when the test is properly performed by a competent operator, the results will be found satisfactory for judging the compliance of the material with the specification. Statements addressing precision and bias are required in ASTM test methods. These then give the user an idea of the precision of the resulting data and its relationship to an accepted reference material or source (if available). Statements addressing determinability are sometimes required as part of the test method procedure in order to provide early warning of a significant degradation of testing quality while processing any series of samples.

5.2 Repeatability and reproducibility are defined in the precision section of every Committee D02 test method. Determinability is defined above in Section 3. The relationship among the three measures of precision can be tabulated in terms of their different sources of variation (see Table 1).

5.2.1 When used, determinability is a mandatory part of the Procedure section. It will allow operators to check their technique for the sequence of operations specified. It also ensures that a result based on the set of determined values is not subject to excessive variability from that source.

5.3 A bias statement furnishes guidelines on the relationship between a set of test results and a related set of accepted

TABLE 1 Sources of Variation

	Method	Apparatus	Operator	Laboratory	Time
Reproducibility	Complete (Result)	Different	Different	Different	Not Specified
Repeatability	Complete (Result)	Same	Same	Same	Almost same
Determinability	Incomplete (Part result)	Same	Same	Same	Almost same

reference values. When the bias of a test method is known, a compensating adjustment can be incorporated in the test method.

5.4 This practice is intended for use by D02 subcommittees in determining precision estimates and bias statements to be used in D02 test methods. Its procedures correspond with ISO 4259 and are the basis for the Committee D02 computer software, *Calculation of Precision Data: Petroleum Test Methods*. The use of this practice replaces that of Research Report RR:D02-1007.[2]

5.5 Standard practices for the calculation of precision have been written by many committees with emphasis on their particular product area. One developed by Committee E11 on Statistics is Practice E691. Practice E691 and this practice differ as outlined in Table 2.

TABLE 2 Differences in Calculation of Precision in Practices D6300 and E691

Element	This Practice	Practice E691
Number of replicates	Two	Any number
Precision is written for	Test method	Each sample
Outlier tests:	Sequential	Simultaneous
Within laboratories	Cochran test	*k*-value
Between laboratories	Hawkins test	*h*-value
Outliers	Rejected, subject to subcommittee approval.	Rejected if many laboratories or for cause such as blunder or not following method.
	Retesting not generally permitted.	Laboratory may retest sample having rejected data.
Analysis of variance	Two-way, applied globally to all the remaining data at once.	One-way, applied to each sample separately.
Precision multiplier	$t\sqrt{2}$, where t is the two-tailed Student's t for 95 % probability.	$2.8 = 1.96\sqrt{2}$
	Increases with decreasing laboratories × samples particularly below 12.	Constant.
Variation of precision with level	Minimized by data transformation. Equations for repeatability and reproducibility are generated in the retransformation process.	User may assess from individual sample precisions.

6. Stages in Planning of an Interlaboratory Test Program for the Determination of the Precision of a Test Method

6.1 The stages in planning an interlaboratory test program are: preparing a draft method of test (see 6.2), planning and executing a pilot program with at least two laboratories (optional but recommended for new test methods) (see 6.3), planning the interlaboratory program (see 6.4), and executing the interlaboratory program (see 6.5). The four stages are described in turn.

6.2 *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 section on precision will be included at this stage only as a heading.

6.2.1 *Interim Repeatability Study*—If the responsible committee decides that an interlaboratory study for the test method is to take place at a later point in time, using this standard, an interim repeatability standard deviation is estimated by following the steps as outlined below. This interim repeatability standard deviation can be used to meet ASTM Form and Style Requirement A21.5.1. When the committee is ready to proceed with the ILS, continue with this practice from 6.3 onwards.

6.2.1.1 *Design*—The following minimum requirements shall be met:

(1) Three (3) samples, compositionally representative of the majority of materials within the design envelope of the test method, covering the low, medium, and high regions of the intended test method range.

(2) Twelve (12) replicates per sample, obtained under repeatability conditions in a single laboratory.

6.2.1.2 *Analysis*—Carry out the following analyses in the order presented:

(1) Perform GESD Outlier Rejection as per Practice D7915 for each sample.

(2) Calculate sample variance (v) and standard deviation (s) for each sample using non-rejected results.

(3) Perform the Hartley test for variance equality as follows:

calculate the ratio : $F_{max} = v_{max}/v_{min}$ where v_{max} and v_{min} are the largest and smallest variance obtained.

(4) If F_{max} is less than 4.85, estimate the interim repeatability standard deviation of the test method by taking the square root of the average variance calculated using individual variances from all samples as illustrated below using three samples:

Interim repeatability standard deviation = $[(v_1 + v_2 + v_3)/3]^{0.5}$, where v_1,v_2, v_3 are variances for each sample; it