Designation: D 3244 – 07a

An American National Standard

Standard Practice for Utilization of Test Data to Determine Conformance with Specifications[1]

This standard is issued under the fixed designation D 3244; 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

The properties of commercial petroleum products are measured by standardized laboratory test methods to check their conformance to specifications. Two or more measurements of the same property of a specific sample by any given test method usually will not give precisely the same answer. Therefore, the test methods generally include a paragraph on the precision of results. This precision is an expression of the reliability of the value of the measured property.

Many difficulties that arise in interpreting specifications are due to test imprecision. Because of this, a true value of a property can never be determined exactly; and it is necessary to infer from measured values the range within which the "true value" is likely to lie. The main purpose of this practice is to indicate how test imprecision should be interpreted relative to specification values.

1. Scope

1.1 This practice covers guidelines and statistical methodologies with which two parties, usually a supplier and a receiver, can compare and combine independently obtained test results to obtain an Assigned Test Value (*ATV*) for the purpose of resolving a product quality dispute.

1.2 This practice defines a technique for comparing an assigned test value with a specification limit.

1.3 This practice applies only to those test methods which specifically state that the repeatability and reproducibility values conform to the definitions herein.

2. Referenced Documents

2.1 *ASTM Standards:* [2]

D 1319 Test Method for Hydrocarbon Types in Liquid Petroleum Products by Fluorescent Indicator Adsorption

D 4057 Practice for Manual Sampling of Petroleum and Petroleum Products

D 4177 Practice for Automatic Sampling of Petroleum and Petroleum Products

D 6300 Practice for Determination of Precision and Bias Data for Use in Test Methods for Petroleum Products and Lubricants

E 29 Practice for Using Significant Digits in Test Data to Determine Conformance with Specifications

2.2 *ISO Standard:*[3]

ISO 4259 Determination and Application of Precision Data in Relation to Methods of Test

3. Terminology

3.1 *Definitions:*

3.1.1 *acceptance limit (AL)*, *n*—a numerical value that defines the point between acceptable and unacceptable quality.

3.1.1.1 *Discussion*—The *AL* is not necessarily the specification limit. It is a value that takes into account the specification value, the test method precision, and the confidence level desired for defining minimum acceptable quality relative to the specification value.

3.1.2 *assigned test value (ATV)*, *n*—the average of all results obtained in the several laboratories which are considered acceptable based on the reproducibility of the test method.

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

3.1.4 *dispute*, *n*—when there is a question as to product quality conformance to specification because a test value obtained falls outside the specification limit(s).

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

Current edition approved Nov. 1, 2007. Published December 2007. Originally published as an appendix to the *1968 Annual Book of ASTM Standards*, Part 18. Originally approved as a standard in 1974. Last previous edition approved in 2007 as D 3244–07.

[2] 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.

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

3.1.5 *operator*, *n*—a person who normally and regularly carries out a particular test.

3.1.6 *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 the repeatability and reproducibility of the test method.

3.1.7 *receiver*, *n*—any individual or organization who receives or accepts the product delivered by the supplier.

3.1.8 *repeatability (r)*, *n*—quantitative expression of the random error associated with a single operator in a given laboratory obtaining replicate results with the same apparatus under constant operating conditions on identical test material within a short period of time. It is defined (3.1.8.1) as that difference between two such single results as would be exceeded in the long run in only 1 case in 20 in the normal and correct operation of the test method (3.1.8.3). (This is known as the 95 % confidence level.)

3.1.8.1 *Discussion*—The repeatability and reproducibility values should have been determined according to the methods described in ASTM Research Report RR:D02-1007, Manual on Determining Precision data for ASTM Methods of Petroleum Products and Lubricants,[4] Practice D 6300, or ISO 4259.

3.1.8.2 *Discussion*—Not all standards organizations define repeatability and reproducibility in precisely these same terms, and attention should always be paid to definitions before comparing precision values quoted.

3.1.8.3 *Discussion*—This difference is related to the repeatability or the reproducibility standard deviation but is not the standard deviation.

3.1.9 *reproducibility (R)*, *n*—quantitative expression of the random error associated with operators working in different laboratories, each obtaining single results on identical test material when applying the same method. It is defined (3.1.8.1) as that difference between two such single and independent results as would be exceeded in the long run in only 1 case in 20 in the normal and correct operation of the test method. See 3.1.8.3.

3.1.10 *result*, *n*—the value obtained by following the complete set of instructions of a test method. It may be obtained from a single determination or several determinations, depending on the instruction of the test method.

3.1.11 *supplier*, *n*—any individual or organization responsible for the quality of a product just before it is taken over by the receiver.

3.1.12 *test sample*, *n*—a portion of the product taken at the place where the product is exchanged, that is, where the responsibility for the product quality passes from the supplier to the receiver. In the event that this is not possible, a suitable sampling location should be mutually agreed upon.

3.1.13 *true value (μ)*, *n*—for practical purposes, the value towards which the average of single results obtained by *N* laboratories using the same standard test method tends, when *N* becomes very large. Consequently, this definition of true value is associated with the particular test method employed.

[4] Supporting data have been filed at ASTM International Headquarters and may be obtained by requesting Research Report RR:D02-1007.

4. Significance and Use

4.1 This practice provides a means whereby the parties to a transaction can resolve potential quality disputes over those product properties which can be tested and expressed numerically.

4.1.1 This practice can be used to ensure that such properties are correctly stated on labels or in other descriptions of the product.

4.1.2 This practice can be implemented in those cases where a supplier uses an in-house or a commercial testing laboratory to sample and test a product prior to releasing the product to a shipper (intermediate receiver) and the ultimate receiver also uses an in-house or commercial testing laboratory to sample and test the product upon arrival at the destination. The assigned test value (*ATV*) would still be determined according to 8.3.

4.2 This practice can assist in the determination of tolerances from specification limits which will ensure that the true value of a property is sufficiently close to the specification value with a mutually agreed probability so that the product is acceptable to the receiver. Such tolerances are bounded by an *acceptance limit* (*AL*). If the *ATV* value determined by applying this practice falls on the *AL* or on the acceptable side of the *AL*, the product can be accepted; otherwise it shall be deemed to have failed the product acceptance requirement established by applying this practice.

4.3 Application of this practice requires the *AL* be determined prior to actual commencement of testing. Therefore, the degree of criticality of the specification, as determined by the Probability of Acceptance (P value) that is required to calculate the *AL*, shall have been mutually agreed upon between both parties prior to execution of actual product testing.

4.3.1 This agreement should include a decision as to whether the *ATV* is to be determined by the absolute or rounding-off method of Practice E 29, as therein defined.

4.3.1.1 If the rounding-off method is to be used, the number of significant digits to be retained must also be agreed upon.

4.3.1.2 These decisions must also be made in the case where only one party is involved, as in the case of a label.

4.3.1.3 In the absence of such an agreement, this practice recommends the *ATV* be rounded in accordance with the rounding-off method in Practice E 29 to the number of significant digits that are specified in the governing specification.

4.4 This practice is designed to be suitable for reference in contracts governing the transfer of petroleum products and lubricants from a supplier to a receiver.

4.5 As a prerequisite for acceptance for lab test results to be used in this practice, the following conditions shall be satisfied:

4.5.1 Long-term standard deviation for the appropriate test method(s) from each lab, as substantiated by in-house quality control programs, on material typical of the product in dispute, shall be statistically equivalent or better than the published method standard deviation under reproducibility conditions.

4.5.2 Each lab shall be able to demonstrate, by way of results from interlaboratory exchange programs, a lack of a systemic bias relative to exchange averages for the appropriate test method(s).

4.5.3 In the event that the long-term standard deviation for any party's laboratory is not statistically equivalent to each other, then, for the purpose of establishing the assigned test value (*ATV*), each laboratory's test result(s) shall be inversely weighted in accordance with laboratory's demonstrated variance(s).

4.6 It is recommended that this practice be conducted under the guidance of a qualified statistician.

5. Sampling

5.1 Sampling should be carried out as specified in accordance with the referenced test method, contract, or specification for the petroleum product under test, Practice D 4057, or Practice D 4177, as appropriate. Obtain enough sample to allow all required determinations to be made. Divide the sample into three secondary samples: a receiver sample, a supplier sample, and a retain sample. The retain sample should itself be large enough to permit further subdivision into three portions in case additional test work is desirable.

6. Applying Test Method Precision Data to Accept or Reject Test Results

6.1 This section describes procedures in which the precision limits of test methods can be used as a decision criterion to accept or reject test results obtained by two laboratories. This section can also be used for acceptance or rejection of results of replicate tests by an operator.

6.2 *Significance of Repeatability (r)*:

6.2.1 *Acceptance of Results*—When only two results are obtained under conditions of repeatability and the difference is equal to or less than the repeatability of the method, the operator may report the average of the two results as being applicable to the sample tested.

6.2.2 *Rejection of Results*—When two results are obtained that differ by more than the repeatability of the method, both should be rejected. Obtain two additional results immediately under conditions of repeatability. If the difference between these two results is equal to or less than the repeatability of the method, the operator should report the average of the two as being applicable to the sample tested. If, however, the difference so obtained again exceeds the repeatability, reject the results and investigate the application of the method.

6.3 *Significance of Reproducibility (R)*:

6.3.1 *Acceptance of Results*—When two results are obtained and comprise one result from each laboratory (Note 1), if the difference is equal to or less than the reproducibility of the method, then both results should be considered acceptable.

NOTE 1—When a comparison for reproducibility is made between results from two laboratories, it is a common practice that single results from each will be compared. If each of the laboratories has produced more than a single result, see 6.4.

6.3.2 *Rejection of Results*—When the results from two laboratories differ by more than the reproducibility of the method, reject both results and each laboratory should repeat the test on the retained sample. If the difference is now equal to or less than the reproducibility, both results should be considered acceptable. If, however, the difference between these results is still greater than the reproducibility, reject the results and investigate the application of the method at each laboratory.

6.4 *Significance of Reduced Reproducibility (R_reduced) from Multiple Testing*—If the number of results obtained in either one or both laboratories is more than one, then the allowable difference between the averages from the two laboratories is given as follows:

$$\text{Difference, } R\_reduced = \sqrt{R^2 - r^2\left(1 - \frac{1}{2n_1} - \frac{1}{2n_2}\right)} \quad (1)$$

where:

R = reproducibility of the method,
r = repeatability of the method,
n_1 = number of results of the first laboratory, and
n_2 = number of results of the second laboratory.

6.5 *Referee Laboratory*—In the event a third or referee laboratory is invited to perform the test using a portion of one of the samples described in 6.3.2, multiply the reproducibility, *R*, by 1.2 (to convert a range for two to a range for three) and compare this value with the difference between the two extreme results for acceptance. If acceptance is indicated, the assigned test value (*ATV*) for the sample should be the average of the three results.

7. Determination of Acceptance Limits by Applying Test Method Precision Data and Specification Criticality Considerations to Specification Limits

7.1 *Specifications*—A specification fixes a limit to the *true value* of a given property. 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 random scattering within tolerances as defined by the test method repeatability and reproducibility limits. Therefore, there is always some uncertainty as to the *true value* of the tested property.

7.2 Although the *true value* is never known exactly, the probability of obtaining any specific test result, relative to a hypothesized true value, can be calculated if the probability distribution function for the test method is known (for example, the Normal or Gaussian distribution).

7.2.1 Some specifications, because of the product characteristic or the end use of the product, or both, require that the receiver have a high degree of assurance that the true value of the product property actually meets or exceeds the quality level indicated by the specification limit value. For the purpose of this practice, such specifications are called *critical* specifications.

7.2.2 Specifications that require assurance only that the product property is not substantially poorer than is indicated by the specification limit are called *noncritical* specifications for the purposes of this practice.

7.3 *Specification Conformance Decision Guidelines*:

7.3.1 Whenever a product is tested for conformity to a specification, a decision must ultimately be made as to whether the product conforms to specification.

7.3.2 The numerical value that delineates the regions of product conformance and non-conformance is the Acceptance Limit (*AL*). The *AL* may or may not coincide with the

specification limit value (*S*) used to define the requirements for the product quality or grade.

NOTE 2—The term "Acceptance" in this context is intended to mean acceptance of the hypothesis that the *true value* of the product property actually meets the quality level indicated by the specification limits. The product may still be accepted or rejected by the receiving party due to other considerations.

7.3.3 The *AL* value, which must be agreed upon by the supplier and receiver prior to commencement of testing, is that level of quality such that, if the *true value* is exactly at the *AL,* either party is willing to take a 50 % chance of either accepting or rejecting the product as tested.

7.3.4 The probability of accepting a product when the true value of the property exactly equals the specification limit value is shown in Fig. 1 and Fig. 2 as a function of $D = (AL - S)/0.255R$, where *D* is a direct measure of the difference between *AL* and *S*. This relationship is based (*1*) on the assumption of normally (Gaussian) distributed testing errors, which is adequate for most test procedures, and (*2*) on using an assigned test value (*ATV*) for making the specification conformance decision that is the average of precision-acceptable results from two laboratories.

7.3.5 The *AL* associated with probability *P* of accepting the product when the true value exactly equals the specification limit value *S* is then given by:

$$AL = S + 0.255 \times R \times D \quad (2)$$

7.3.5.1 The factor 2.55 in Eq 2 is for *N* (no. of labs) = 2. For *N* greater than 2, the 0.255 factor should be multiplied by $\sqrt{2/N}$.

7.3.6 In the absence of an agreement to the contrary, this practice recommends that for non-critical specifications, the *AL* is set such that there is 95 % probability that the product will be accepted if the true value of the property is exactly at the specification limit value. Thus, the *AL* will be set by using a confidence level $P = 0.95$ as shown in 7.3.5. It should be noted that for $P = 0.95$, the *AL* will actually be numerically outside the specification limit values.

7.3.7 In the absence of an agreement to the contrary, this practice recommends that for critical specifications, the *AL* is set such that there is 5 % probability that the product will be accepted if the true value of the property is exactly at the specification limit value. Thus, the *AL* will be set by using a confidence level $P = 0.05$ as shown in 7.3.5. It should be noted that for $P = 0.95$, the *AL* will actually be numerically inside the specification limit values.

7.3.8 When $D = 0$, the *AL* coincides exactly with the specification limit. The *P* value for $D = 0$ is 0.5, which means that there is a 50 % probability that the product will be accepted if the true value of the property is exactly at the specification limit. This is also the delineation point between critical and non-critical specification as chosen by this practice.

7.3.8.1 For specifications having both minimum and maximum limits, the procedure in 7.3.5 must be applied twice to give both upper and lower *AL*s. There must be some allowable region remaining between the lower and upper *AL*s.

7.3.9 When only a single test result is or will be available, the relationships given should be used with $N = 1$ (7.3.5.1). Obviously, no check on reproducibility precision can be made with a single test result, and the single value becomes the *ATV* for the sample.

7.3.10 The relationships between the *AL*s for critical and noncritical specifications are shown in Fig. 3 for a minimum specification.

			$D = (AL - S)/0.255\,R$	
		Probability (*P*) of Acceptance	Maximum Specification Limit	Minimum Specification Limit
Critical Spec Region		0.001	-3.090	3.090
		0.005	-2.576	2.576
		0.010	-2.326	2.326
		0.025	-1.960	1.960
	Recommended P=>	0.050	-1.645	1.645
		0.100	-1.282	1.282
		0.150	-1.036	1.036
		0.200	-0.842	0.842
		0.300	-0.524	0.524
Noncritical Spec Region		0.500	0.000	0.000
		0.700	0.524	-0.524
		0.800	0.842	-0.842
		0.850	1.036	-1.036
		0.900	1.282	-1.282
	Recommended P =>	0.950	1.645	-1.645
		0.975	1.960	-1.960
		0.990	2.326	-2.326
		0.995	2.576	-2.576
		0.999	3.090	-3.090

NOTE—Based on $N = 2$ = number of different laboratories' results used to obtain *ATV*. See text for use of this table.

FIG. 1 Deviation of *AL* from Specification for Product Acceptance at a Given Probability

8. Obtaining the Assigned Test Value (*ATV*)

8.1 The following procedure will produce an *ATV* with precision control based on the reproducibility of the test method.

8.2 The receiver and supplier should obtain independent test results, X_R and X_S, respectively.

NOTE 3—The supplier's result must be on the *test sample* (see Section 5) and not a reported value by the supplier. In many cases, a reported value by the supplier is obtained on a different sample, for example, at point of manufacture, and may be the average of several determinations.

8.3 *ATV Procedure*:

8.3.1 If the absolute value of $\Delta = X_R - X_S$ is less than or equal to *R*, the reproducibility of the test method, average the two results to obtain the following in accordance with 6.3.1:

$$ATV = (X_R + X_S)/2 \quad (3)$$

8.3.2 If the absolute value of Δ exceeds *R*, reject both results and retest on portions of the retain sample to obtain X_R', X_S'.

8.3.3 If the absolute value of $\Delta' = X_R' - X_S'$ is less than or equal to *R*, average the two results to obtain the following in accordance with 6.3.2:

$$ATV = (X_R' + X_S')/2 \quad (4)$$

8.3.4 If the absolute value of Δ′ exceeds *R*, obtain a new test value X_{RL} from a referee laboratory (6.5).

8.3.5 If $\Delta_3 = X_{max} - X_{min}$ is less than or equal to 1.2 *R*, obtain the following:

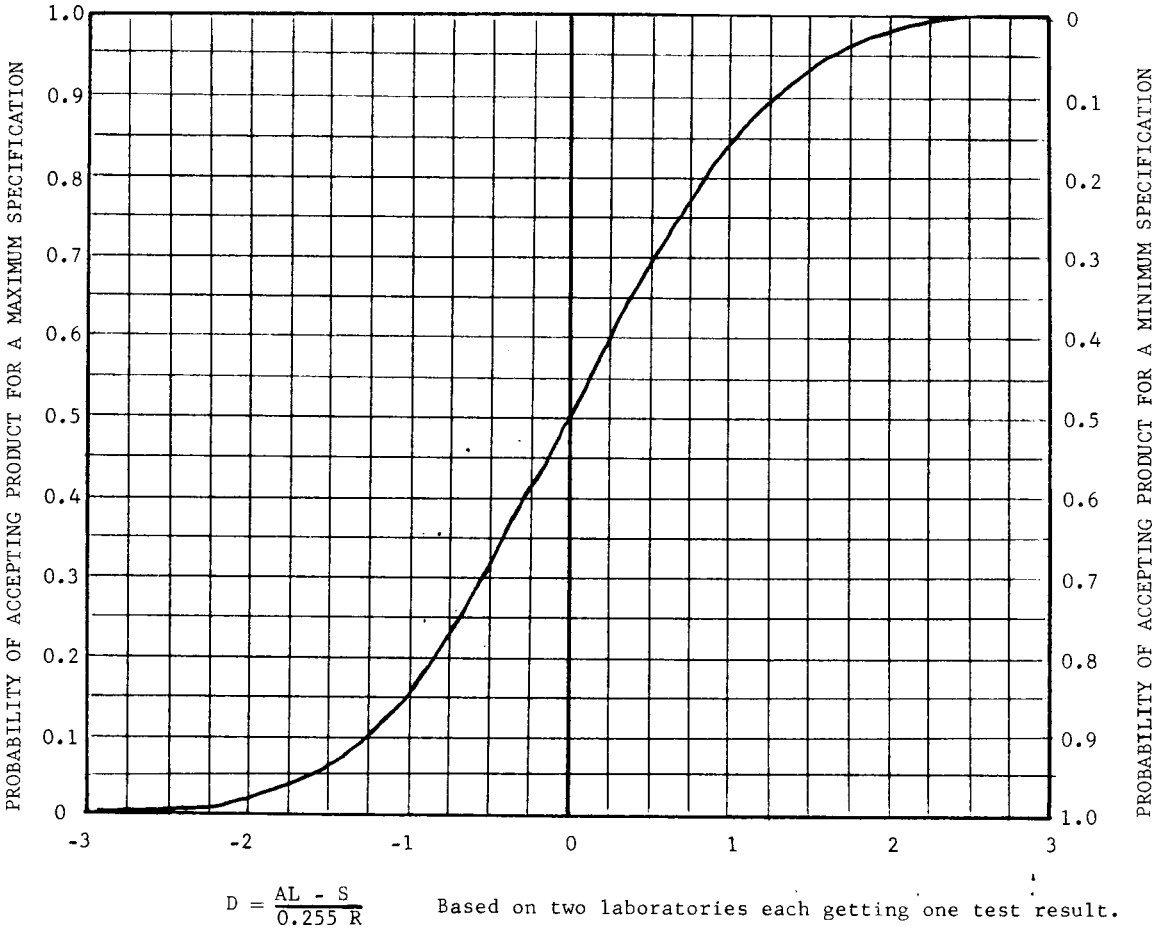
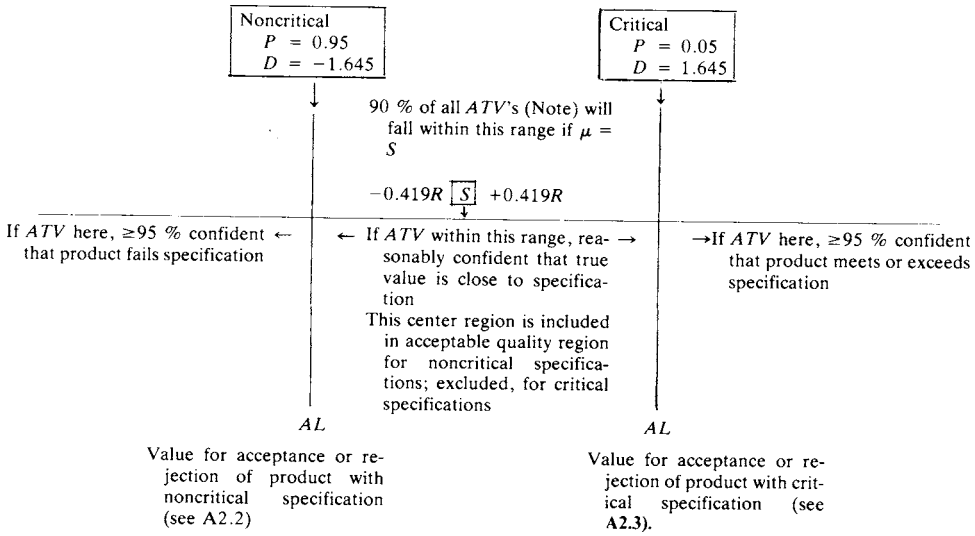


FIG. 2 Probability of Acceptance vs Deviation of *AL* from True Value = S



NOTE—This applies when *ATV* is established by the average of two results, one each from two different laboratories.

FIG. 3 Relationships Between *AL* s for Critical and Noncritical Specifications

$$ATV = (X_R' + X_S' + X_{RL})/3 \quad (5)$$

8.3.6 If Δ_3 exceeds 1.2 *R*, obtain *ATV* as the average of the closer pair.

NOTE 4—This last step for obtaining an *ATV* does not comply rigidly to statistical concepts. It is done in this manner because in most cases the test sample (see Section 5) is depleted.

8.4 The above procedure will always yield an *ATV*. If the supplier's and receiver's laboratories have little or no bias relative to each other, then the procedure will end at 8.3.1 about 95 % of the time, and some 95 % of the remaining 5 %, at 8.3.3.

8.5 If any particular supplier and receiver pair find they frequently must go as far as calling for a reference laboratory test, they should carefully check their running of the test, as well as examine their calibration practice versus other laboratories that have demonstrated proficieny in the conduct of the particular test method.

8.6 This procedure for obtaining an *ATV* is designed for the test of samples obtained according to Section 5.

8.6.1 If more extensive testing is needed for special situations, comparable procedures can be developed. A statistician or quality control expert should be consulted to do this.

9. Product Quality Conformance

9.1 A product should be considered as conforming to the specifications if the *ATV* of each property meets the *AL* value.

9.2 The supplier should ship product only if there is confidence that each property meets specification values.

9.3 When the receiver has obtained a single result, the product quality should be considered suspect if the test result fails the *AL* value (see A3.1.5).

9.4 A dispute between supplier and receiver may arise whenever a receiver's result fails the *AL* value.

9.5 The dispute should be resolved by obtaining an assigned test value (*ATV*) for the product as an estimate of the "true value" and comparing this to the acceptance limit (*AL*) as determined in 7.3.

10. Acceptance or Rejection of Product

10.1 If the *ATV* is equal to or better than the *AL* limit, the product is to be accepted as having met specification.

10.2 If the *ATV* fails the *AL* value, the product is to be rejected as failing specification.

10.3 These concepts are presented graphically in Fig. 4.

10.3.1 The plotted lines are boundary conditions separating acceptable results from those indicating other alternative actions.

10.3.1.1 The sample is considered acceptable if the two results fall to the left of the line, $(X_R + X_S)/2 = ATV = AL$ if they are also within the lines. $X_R - X_S = \pm R$.

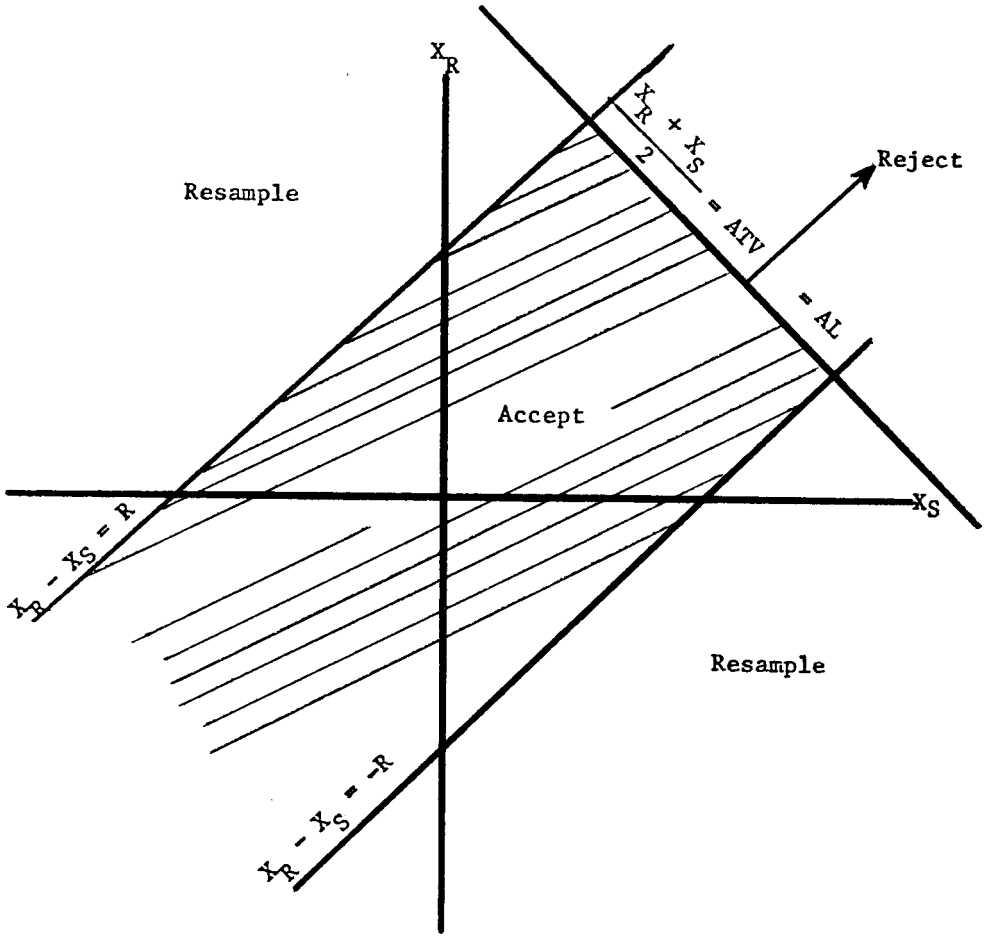


FIG. 4 Diagram Showing Regions of Acceptance, Rejection, and Resampling

10.3.2 The sample is unacceptable if the results lie to the right of the line $(X_R + X_S)/2 = ATV = AL$.

10.3.3 Initial results falling in the region labeled *resample* call for a retest.

10.3.3.1 If results for a second sample also fall in the resample region, a reference laboratory should be included in the new testing program.

10.4 The actual consequences of rejecting a product for failure to meet specification are subject to prior agreement or negotiation between the parties concerned.

11. Keywords

11.1 acceptance; acceptance limits; agreement; conformance; dispute; precision; rejection; specifications

ANNEXES

(Mandatory Information)

A1. GUIDES FOR DETERMINING *AL*

A1.1 As *AL* is the dividing line between acceptable and unacceptable test results, it is an important step in determining conformance to specification.

A1.2 The probability of rejection or acceptance of any product whose *true value* is *AL* is always 50 %, regardless of the precision of the *ATV* value used in making the decision. This statement requires only the assumption of a symmetric distribution of testing errors (such as, but not restricted to, the normal distribution).

A1.3 Referring to 7.3.9, to determine an *AL* that will give a desired probability *P* that the product is accepted: For noncritical specifications, the *P* value may be chosen to be fairly large, perhaps 0.90 or 0.95; for critical specifications, *P* would be chosen below 0.50, perhaps 0.05 or 0.10. Even lower values would be called for in cases of extreme criticality.

A1.4 For critical specifications, the product is acceptable only if the *ATV* is better than *S* at nearly the 100 $(1 - P)$ % significance level.

A1.5 For noncritical specifications, the product is rejected only if the *ATV* is worse than *S* at nearly the 100 *P* % significance level.

A2. EXAMPLES OF *AL* DETERMINATION AND USE

A2.1 Assume that we are testing a product whose quality is measured by ASTM D XYZ which has a repeatability of 1 and a reproducibility of 2. There is a maximum specification of 10.0 for the property measured by D XYZ. In any case, the supplier will not ship the product unless his tests at point of manufacture show that the limit of 10 has not been exceeded. Only two laboratories, supplier's and receiver's, will make tests to determine *ATV* ($N = 2$).

A2.2 *Noncritical Specification*—Receiver establishes a limit of 10 maximum as a noncritical specification with $P = 0.95$.

A2.2.1 At $P = 0.95$, from Fig. 1 or Fig. 2, obtain $D = 1.645$.

A2.2.2 $AL = S + 0.255 \quad R \cdot D$ (from 7.3.5). $AL = 10 + 0.255 \times 2 \times 1.645 = 10.84$. Product as tested must average (*ATV*) 10.84 or lower to be acceptable.

A2.2.3 Upon testing of sample (Section 8), receiver obtains test result $X_R = 10.8$, supplier obtains $X_S = 9.9$. $\Delta = 10.8 - 9.9 = 0.9 < R = 2$, meeting the reproducibility requirement, so that

$$ATV = (10.8 + 9.9)/2 = 10.34 \qquad \text{(A2.1)}$$

A2.2.4 The *ATV* as obtained is below *AL*, so the product is accepted.

A2.3 *Critical Specification*—Another receiver needs, for reasons known best to him, a very high level of assurance that the product meets the specification of 10.0.

A2.3.1 Using $P = 0.025$, obtain $D = -1.960$ from Fig. 1 (Fig. 2).

A2.3.2 $AL = S + 0.255 \quad R \cdot D$. $AL = 10 + 0.255 \times 2 \times (-1.960) = 9.00$. Thus, the product as tested (*ATV*) must average 9.00 or lower to be acceptable.

A2.3.3 Testing of sample (Section 8) gives

$$X_R = 9.4 \qquad \text{(A2.2)}$$
$$X_S = 9.2$$

$$\Delta = 0.2 \text{ meets reproducibility criterion} \qquad \text{(A2.3)}$$

thus $ATV = (9.4 + 9.2)/2 = 9.3$.

A2.3.4 The *ATV* as obtained is above the *AL* so the product is rejected as unacceptable in quality even though the *ATV* is better than the specification value.

A2.4 *Converting a Critical Specification to a Noncritical Specification:*

A2.4.1 The receiver in the example of A2.3 wanted a high level of assurance that the product met a specification of 10 and, hence, applied a low value for *P* in establishing the *AL*. He could have used a noncritical specification of 8.16 to accomplish the same objective.

A2.4.2 To obtain a noncritical specification value having the same *AL* as a critical specification value, solve the equation of 7.3.5.

$$AL = S + (0.255)(R)(D) \quad \text{(A2.4)}$$

for S using the value of 9.00 from A2.3.2. A value of $D = 1.645$ is used for a noncritical specification. Thus,

$$S = 9.00 - (0.255)(2)(1.645) = 8.16 \quad \text{(A2.5)}$$

A2.4.3 In reality, the actual quality of product needed by this receiver is 1.84 units better (10 – 8.16) than the receiver described in A2.1.

A3. CONSTANTS USED IN EQUATIONS

A3.1 The constant used in equation of 7.3.5 was developed as shown in the following paragraphs.

A3.1.1 The *AL* equals the specification value plus a term reflecting the probability of a difference between the true value = *S* and an observed value of a given property. Thus,

$$AL = S + (\sigma D/\sqrt{N}) \quad \text{(A3.1)}$$

where:

- σ = standard deviation of measurement for the test method under reproducibility conditions,
- D = deviation between a measured value and the true value for a specified probability, and
- N = number of different laboratories whose test results are averaged to establish the assigned test value (*ATV*).

A3.1.2 The definition for reproducibility (3.1.9) is as follows:

$$R = \sigma t_{95}\sqrt{2} \quad \text{(A3.2)}$$

where the t_{95} value is 1.96 for 95 % probability or confidence level. Thus,

$$R = \sigma\,(1.96)(\sqrt{2}) = 2.77\,\sigma \quad \text{(A3.3)}$$

$$\text{or} \quad \sigma = R/2.77 = 0.361R \quad \text{(A3.4)}$$

A3.1.3 When the assigned test value (*ATV*) is obtained by averaging two results from two different laboratories, $N = 2$. Substituting the values of $\sigma = 0.361R$ from A3.1.2 and $N = 2$ in the equation of A3.1.1 gives the following:

$$AL = S + (0.361RD)\sqrt{2} = S + 0.255\,RD \quad \text{(A3.5)}$$

A3.1.4 For the conditions described in 7.3.6, the value for *D* from Fig. 1 for 95 % assurance of acceptance of a product meeting specification exactly is + 1.645 for a maximum specification and – 1.645 for a minimum specification substituting these values in the equation of A3.1.3 gives the following for a *maximum* specification:

$$AL = S + (0.255)(+1.645)(R) = S + 0.419\,R \quad \text{(A3.6)}$$

and for a *minimum* specification:

$$AL = S + (0.255)(-1.645)(R) = S - 0.419\,R \quad \text{(A3.7)}$$

The constant, 0.419, is the one shown in the diagram of 7.3.10.

A3.1.5 It is emphasized that the constants developed for calculation of *AL* are based on an *ATV* established by averaging two test results, one each from two different laboratories. If a result from only one laboratory is used to determine the AQL, the value for *N* is one and the equation for establishing the *AL* in accordance with A3.1.3 is as follows:

$$AL = S + \frac{(0.361R)(D)}{\sqrt{1}} = S + 0.361\,RD \quad \text{(A3.8)}$$

and the equation for a *maximum* specification in accordance with A3.1.4 becomes

$$AL = S + (0.361)(1.645)R = S + 0.594\,R \quad \text{(A3.9)}$$

and for a *minimum* specification

$$AL = S - 0.594\,R \quad \text{(A3.10)}$$

A3.1.6 The equations presented in A3.1.5 should be used to establish *AL* for comparison with a result from a single laboratory. A single result is usually insufficient to estimate the *ATV* of a property with a high degree of accuracy. If the single observed value does not meet the *AL* in accordance with A3.1.5, additional testing and investigation is justified.

A4. EXAMPLES OF LABORATORY PROFICIENCY PREREQUISITES

A4.1 The data in Table A4.1 will be used to demonstrate the proficiency prerequisites of 4.5.2 and 4.5.3. (The method in this example is Test Method D 1319 - Volume Percent Saturates.) In this case, Laboratories A, B, and C have all participated in the same interlaboratory exchange program, and there have been six samples exchanged.

A4.2 The deviations from the respective sample averages are computed for every result, and the average and standard deviation of the deviations are computed for each laboratory. See Table A4.2.

A4.3 In order to test for a statistically significant laboratory bias, as required in 4.5.2, a *t*-test may be conducted for each laboratory as follows:

A4.3.1 Compute the standard error of the deviations for the laboratory by dividing the standard deviation by the square root of the number of exchanges in which the laboratory has participated. These are shown in Table A4.2.

A4.3.2 The *t*-statistic is the ratio of the average deviation divided by the standard error for the laboratory. The *t*-statistics for the xample are in Table A4.2.

A4.3.3 The degrees of freedom are one less than the number of exchanges in which the laboratory has participated. In the example, there are five degrees of freedom for every laboratory.

A4.3.4 If the absolute value of the *t*-statistic exceeds the 95th percentile of Student's |*t*| distribution with the appropriate degrees of freedom, then a statistically significant laboratory bias exists. The appropriate |*t*| percentiles are found in Table A4.3. As the absolute *t*-statistic for Laboratory C, 2.71, exceeds the |*t*| percentile with 5 degrees of freedom, 2.57, results from this laboratory may not be used in determining an assigned test value (*ATV*).

A4.4 Long-term standard deviations for each laboratory are estimated from in-house quality control programs. In the event that these data are not comparable, then data from interlaboratory exchange programs may be used to test for statistical equivalence, as required in 4.5.3. An *F*-test may be conducted, as follows:

A4.4.1 The *F*-ratio for comparing two laboratories, standard deviations of deviations is formed by dividing the square of the larger standard deviation by the square of the smaller. In our example, comparing Laboratories A and B, $F = 4.88^2/1.33^2 = 23.8/1.77 = 13.5$.

A4.4.2 If the *F*-ratio exceeds the 95th percentile of the *F*-distribution with the appropriate number of degrees of freedom for the numerator and for the denominator, then the two laboratories' standard deviations are not statistically equivalent. The appropriate *F* percentiles are found in Table A4.4, and as the *F*-ratio of 13.5 exceeds the 95th percentile for 5 and 5 degrees of freedom, 7.15, we conclude that Laboratory A and Laboratory B have different long-term standard deviations.

A4.5 In the event that two laboratories with long term standard deviations that are not equivalent are to obtain an *ATV*, then each laboratory's result shall be inversely weighted by the laboratory's demonstrated variance. For example, if Laboratory A obtained a single result of 51.1, while Laboratory B obtained 47.8, then the *ATV* would be computed as

$$(51.1/1.33^2+47.8/4.88^2)/(1/1.33^2 + 1/4.88^2) = 50.9. \quad \text{(A4.1)}$$

TABLE A4.1 Sample Data from Exchange Program

Laboratory	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
A	53.3	61.6	54.8	44.9	57.2	62.9
B	56	61.9	52.7	39.6	57	50
C	30.9	50.8	58.5	35.1	50.4	38.2
(more labs)	. . .	. . .	. . .	. . .	. . .	. . .
No. of Laboratories	47	56	60	48	67	45
Mean	53.8	59.8	55.5	44.5	56.1	60.2

TABLE A4.2 Deviations from Exchange Means

Laboratory	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Mean	Standard Deviation	Standard Error	*t*	Degrees of Freedom
A	-0.5	1.8	-0.7	0.4	1.1	2.7	0.8	1.33	0.54	1.48	5
B	2.2	2.1	-2.8	-4.9	0.9	-10.2	-2.1	4.88	1.99	-1.06	5
C	-22.9	-9	3	-9.4	-5.7	-22	-11	9.93	4.05	-2.71	5

TABLE A4.3 95th Percentiles of the Distribution of $|t|$

Degrees of Freedom	t
4	2.78
5	2.57
6	2.45
7	2.36
8	2.31
9	2.26
10	2.23
12	2.18
15	2.13
20	2.09
25	2.06
30	2.04
60	2.00
120	1.98

TABLE A4.4 95th Percentiles of the *F* Distribution

NOTE 1—When F is defined as the ratio of the larger to the smaller mean square.

Denominator Degrees of Freedom	Numerator, Degrees of Freedom													
	4	5	6	7	8	9	10	12	15	20	25	30	60	120
4	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.50	8.46	8.36	8.31
5	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.27	6.23	6.12	6.07
6	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.11	5.07	4.96	4.90
7	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.40	4.36	4.25	4.20
8	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00	3.94	3.89	3.78	3.73
9	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.60	3.56	3.45	3.39
10	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.35	3.31	3.20	3.14
12	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.01	2.96	2.85	2.79
15	3.80	3.52	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.69	2.64	2.52	2.46
20	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.40	2.35	2.22	2.16
25	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.30	2.23	2.18	2.05	1.98
30	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.20	2.12	2.07	1.94	1.87
60	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.87	1.82	1.67	1.58
120	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94	1.82	1.75	1.69	1.53	1.43