Designation: D 2906 – 97

Standard Practice for Statements on Precision and Bias for Textiles[1]

This standard is issued under the fixed designation D 2906; 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

Work was begun in August 1966 on recommendations for statements on precision and accuracy. The first recommendations were issued as ASTMD-13 White Paper, Statements on Precision and Accuracy, MARK I, December 1968, prepared by Subcommittee C-6 on Editorial Policy and Review. After a decision that the recommendations should be a recommended practice under the responsibility of Subcommittee D13.93, Sampling, Presentation and Interpretation of Data, the recommendations were revised and published as Practice D 2906 – 70 T.

Information was added in Practice D 2906 – 73 on methods (*1*) for which precision has not been established, (*2*) for which test results are not variables, and (*3*) for which statements are based on another method. Practice D 2906 – 74 was expanded to include test methods in which test results are based on the number of successes or failures in a specified number of observations or on the number of defects or instances counted in a specified interval of time or in a specified amount of material. The present text provides for a nontechnical summary at the beginning of recommended texts based on normal distributions or on transformed data and for a more positive statement on accuracy when the true value of a property can be defined only in terms of a test method.

In 1984, changes were introduced to replace the term "accuracy" with "bias" as directed in the May 1983 edition of *Form and Style for ASTM Standards.*

1. Scope

1.1 This practice serves as a guide for using the information obtained as directed in Practice D 2904 or obtained by other statistical techniques from other distributions, to prepare statements on precision and bias in ASTM methods prepared by Committee D-13. The manual on form and style for standards specifies that statements on precision and bias be included in test methods.[2] Committee D-13 recommends at least a statement about single-operator precision in any new test method, or any test method not containing a precision statement that is put forward for 5-year approval, in both instances with a complete statement at the next reapproval. If a provisional test method is proposed, at least a statement on single-operator precision is expected.

1.2 The preparation of statements on precision and bias requires a general knowledge of statistical principles including the use of components of variance estimated from an analysis of variance. Instructions covering such calculations are available in Practice D 2904 or in any standard text **(1,2,3,4,** and **5).**[3]

1.3 The instructions in this practice are specifically applicable to test methods in which test results are based (*1*) on the measurement of variables, (*2*) on the number of successes or failures in the specified number of observations, (*3*) on the number of defects or incidents counted in a specified interval or in a specified amount of material, and (*4*) on the presence or absence of an attribute in a test result (a go, no-go test). Instructions are also included for methods of test for which precision has not yet been estimated or for which precision and accuracy have been reported in another method of test. For observations based on the measurement of variables, the instructions of this practice specifically apply to test results that are the arithmetic average of individual observations. With qualified assistance, the same general principles can be applied to test results that are based on other functions of the data such as standard deviations.

[1] This practice is under the jurisdiction of ASTM Committee D-13 on Textiles and is the direct responsibility of Subcommittee D13.93 on Statistics.

Current edition approved Nov. 10, 1997. Published August 1998 . Originally published as D 2906 – 70 T. Last previous edition D 2906 – 91.

[2] *Form and Style for ASTM Standards*, May 1983, available from American Society for Testing and Materials, 100 Barr Harbor Drive, West Conshohocken, PA 19428.

[3] The boldface numbers in parentheses refer to the list of references at the end of this practice.

1.4 This standard includes the following sections:

2. Referenced Documents

2.1 *ASTM Standards:*

D 123 Terminology Relating to Textiles[4]

D 2904 Practice for Interlaboratory Testing of a Textile Test Method that Produces Normally Distributed Data[4]

D 2905 Practice for Statements on Number of Specimens for Textiles[4,5]

E 456 Terminology Relating to Quality and Statistics[6]

E 691 Practice for Conducting An Interlaboratory Study to Determine the Precision of a Test Method[6]

2.2 ASTM Adjuncts:

TEX-PAC[7]

NOTE 1—Tex-Pac is a group of PC programs on floppy disks, available through ASTM Headquarters, 100 Barr Harbor Drive, West Conshohocken, PA 19428, USA. The calculations of critical differences and confidence limits described in the various sections of this practice can be performed using some of the programs in this adjunct.

3. Terminology

3.1 *Definitions:*

3.1.1 *accuracy*, *n—of a test method*, the degree of agreement between the true value of the property being tested (or accepted standard value) and the average of many observations made according to the test method, preferably by many observers. See also *bias* and *precision*.

3.1.1.1 *Discussion*—Increased accuracy for a test method is associated with decreased bias relative to an accepted reference value. Although the total bias of a test method is equivalent to the accuracy of the test method, the present edition of *Form and Style for ASTM Standards* recommends using the term "bias" since the accuracy of individual observed values is sometimes defined as involving both the precision and the bias of the method.

3.1.2 *bias*, *n—in statistics*, a constant or systematic error in test results.

3.1.2.1 *Discussion*—Bias can exist between the true value and a test result obtained from one method, between test results obtained from two methods, or between two test results obtained from a single method, for example, between operators or between laboratories.

3.1.3 *characteristic*, *n*—a property of items in a sample or population which, when measured, counted, or otherwise observed, helps to distinguish between the items. (E 456)

3.1.4 *confidence level*, *n*—the stated proportion of times the confidence interval is expected to include the population parameter. (E 206)

3.1.4.1 *Discussion*—Statisticians generally accept that, in the absence of special considerations, 0.95 or 95 % is a realistic confidence level. If the consequences of not including the unknown parameter in the confidence interval would be grave, then a higher confidence level might be considered which would lengthen the reported confidence interval. If the consequences of not including the unknown parameter in the confidence interval are of less than usual concern, then a lower confidence level might be considered which would shorten the reported confidence interval.

3.1.5 *critical difference*, *n*—the observed difference between two test results which should be considered significant at the specified probability level.

3.1.5.1 *Discussion*—The critical difference is not equal to the expected variation in a large number of averages of observed values; it is limited to the expected difference between only two such averages and is based on the standard error for the difference between two averages and not on the standard error of single averages.

3.1.6 *laboratory sample*, *n*—a portion of material taken to represent the lot sample, or the original material, and used in the laboratory as a source of test specimens.

3.1.7 *lot sample*, *n*—one or more shipping units taken at random to represent an acceptance sampling lot and used as a source of laboratory samples.

3.1.8 *parameter*, *n—in statistics*, a variable that describes a characteristic of a population or mathematical model.

3.1.9 *percentage point*, *n*—a difference of 1 percent of a base quantity.

3.1.9.1 *Discussion*—A phrase such as "a difference of X %" is ambiguous when referring to a difference in percentages. For example, a change in the moisture regain of a material from 5 % to 7 % could be reported as an increase of 40 % of the initial moisture regain or as an increase of two percentage points. The latter wording is recommended.

3.1.10 *precision*, *n*—the degree of agreement within a set of observations or test results obtained as directed in a method.

3.1.10.1 *Discussion*—The term "precision", delimited in various ways, is used to describe different aspects of precision. This usage was chosen in preference to the use of "repeatability" and "reproducibility" which have been assigned conflicting meanings by various authors and standardizing bodies.

3.1.11 *precision*, *n—under conditions of single-operator precision*, the single-operator-laboratory-sample-apparatus-day precision of a method; the precision of a set of statistically independent observations all obtained as directed in the method and obtained over the shortest practical time interval in one laboratory by a single operator using one apparatus and randomly drawn specimens from one sample of the material being tested.

[4] *Annual Book of ASTM Standards*, Vol 07.01.
[5] *Annual Book of ASTM Standards*, Vol 07.02.
[6] *Annual Book of ASTM Standards*, Vol 14.02.
[7] PC programs on floppy disks are available through ASTM. Request ADJD2906.

3.1.11.1 *Discussion*—Results obtained under conditions of single-operator precision represent the optimum precision that can be expected when using a method. Results obtained under conditions of within-laboratory and between-laboratory precision represent the expected precision for successive test results when a method is used respectively in one laboratory and in more than one laboratory.

3.1.12 *precision, n—under conditions of within-laboratory precision with multiple operators*, the multi-operator, single-laboratory-sample, single-apparatus-day (within operator) precision of a method; the precision of a set of statistically independent test results all obtained in one laboratory using a single sample of material and with each test result obtained by a different operator, with each operator using one apparatus to obtain the same number of observations by testing randomly drawn specimens over the shortest practical time interval.

3.1.13 *precision, n—under conditions of between-laboratory precision*, the multi-laboratory, single-sample, single-operator-apparatus-day (within-laboratory) precision of a method; the precision of a set of statistically independent test results all of which are obtained by testing the same sample of material and each of which is obtained in a different laboratory by one operator using one apparatus to obtain the same number of observations by testing randomly drawn specimens over the shortest practical time interval.

3.1.14 *probability level, n*—a general term that reflects the stated proportion of times an event is likely to occur. (Compare to *confidence level* and *significance level*.)

3.1.15 *sample, n*—(*1*) a portion of material which is taken for testing or for record purposes. (See also *sample lot; sample, laboratory; and specimen.*) (*2*) a group of specimens used, or of observations made, which provide information that can be used for making statistical inferences about the population(s) from which the specimens are drawn.

3.1.16 *significance level, (α), n*—the stated upper limit for the probability of a decision being made that an hypothesis about the value of a parameter is false when in fact it is true.

3.1.17 *specimen, n*—a specific portion of a material or a laboratory sample upon which a test is performed or which is selected for that purpose. (*Syn.* test specimen.)

3.1.18 *test result, n*—a value obtained by applying a given test method, expressed as a single determination or a specified combination of a number of determinations.

3.1.19 For definitions of other textile terms used in this practice, refer to Terminology D 123. For definitions of other statistical terms used in this practice, refer to Terminology D 123 or Terminology E 456.

4. Statements on Acceptance Testing

4.1 In the section on *Significance and Use*, include a statement on the use of the method for acceptance testing. If the determined precision supports such use, the test method should be recommended for use. If not, the test method should not be recommended for use. Other circumstances may cause a test method to be used for acceptance testing when precision is poor, or precision is not known. In an event such may occur, advice may be given on the consequences of such usage. In most cases, one of the recommended texts in 4.1.1, 4.1.3, or 4.1.5 will be adequate. In all cases, the recommended text in 4.2 may be part of the statement.

NOTE 2—The final decision to use a specific method for acceptance testing of commercial shipments must be made by the purchaser and the supplier and will depend on considerations other than the precision of the method, including the cost of sampling and testing and the value of the lot of material being tested.

4.1.1 If serious disagreements between laboratories is relatively unlikely, consider the following statement (Note 3).

NOTE 3—In these recommended texts, the numbers of sections, notes, footnotes, equations, and tables are for illustrative purposes and are not intended to conform to the numbers of other parts of this practice. In correspondence they can be best referenced by such phrases as: "the illustrative text in 4.1.1 numbered as 4.1.2."

4.1.2 Method D 0000 for the determination of (insert here the name of the property) is considered satisfactory for acceptance testing of commercial shipments of (insert here the name of the material) since (insert here the specific reason or reasons, such as: (*1*) current estimates of between-laboratory precision are acceptable, (*2*) the method has been used extensively in the trade for acceptance testing, or (*3*) both of the preceding reasons.)

4.1.3 If it is relatively likely that serious disagreements between laboratories may occur but the method is the best available, consider the following statement (Note 3).

4.1.4 Method D 0000 for the determination of (insert here the name of the property) may be used for the acceptance testing of commercial shipments of (insert here the name of the material) but caution is advised since (insert here the specific reason or reasons, such as: (*1*) information on between-laboratory precision is lacking or incomplete or (*2*) between-laboratory precision is known to be poor.) Comparative tests as directed in 4.2.1 may be desirable.

4.1.5 If a method is not recommended for acceptance testing because a more appropriate method is available, because the test is intended for development work only, or because experience has shown that results in one laboratory cannot usually be verified in another laboratory, consider the following statement.

4.1.6 Method D 0000 for the determination of (insert here the name of the property) is not recommended for the acceptance testing of commercial shipments of (insert here the name of the material) since (insert here the specific reason or reasons, such as: (*1*) an alternative, Method D 0000 is recommended for this purpose because (insert here reasons such as those in the illustrative text following 4.1.1), (*2*) experience has shown that results in one laboratory cannot usually be verified in another laboratory, or (*3*) the scope of the method states that the method is recommended only for development work within a single laboratory). Although Method D 0000 is not recommended for use in acceptance testing, it is useful because (insert here the reason or reasons the subcommittee thinks the method should be included in the *Annual Book of ASTM Standards*).

4.2 Include the following statement on conducting comparative tests between laboratories as part of all statements on the use of a method for acceptance testing.

4.2.1 In case of a dispute arising from differences in reported test results when using Method D 0000 for acceptance

testing of commercial shipments, the purchaser and the supplier should conduct comparative tests to determine if there is a statistical bias between their laboratories. Competent statistical assistance is recommended for the investigation of bias. As a minimum, the two parties should take a group of test specimens that are as homogeneous as possible and that are from a lot of material of the type in question. The test specimens should then be randomly assigned in equal numbers to each laboratory for testing. The average results from the two laboratories should be compared using Student's *t*-test for unpaired data and an acceptable probability level chosen by the two parties before the testing is begun. If a bias is found, either its cause must be found and corrected or the purchaser and the supplier must agree to interpret future test results in the light of the known bias.

NOTE 4—The test of significance specified in the illustrative text for 4.2.1 is appropriate only for unpaired data from normal distributions. If the type of distribution is not known or is known not to be normal, substitute "a nonparametric test for unpaired data" for "Student's *t*-test for unpaired data" in the next to last sentence in 4.2.1. If the same specimens are evaluated in both laboratories, the description of the preparation of specimens will need to be altered and either "Student's *t*-test for paired data" or "a nonparametric test for paired data" used to describe the test of significance in the next to last sentence of the proposed text for 4.2.1.

5. Statistical Data in Two Sections of Methods

5.1 Many methods approved by Committee D-13 include a section on "*Number of Specimens*" which normally does not describe any interlaboratory testing done during the development of the method or include any estimates of the components of variance obtained from such a study. When that section is written as directed in Practice D 2905, the text consists of three parts. The first part specifies the allowable variation, the probability level, and whether one-sided or two sided limits are required. The second part specifies how the number of observations required for the desired precision can be calculated from an estimate of the single-operator component of variance based on records of the specific laboratory involved. The third part specifies a definite number of observations to be made in the absence of adequate information about the single-operator precision. In the last case, the recommended number of observations is based on a value of the single-operator component of variance somewhat larger than is usually found in practice. Thus, the inexperienced user has the protection of a relatively large number of observations. However, the section on *Number of Specimens* does not allow the inexperienced user of the method to visualize the single-operator precision of the method to be expected for averages based on different numbers of specimens tested by experienced operators. The desirability of supplying such information is the primary reason for requiring information about single-operator precision even when the section on *Number of Specimens* is based on Practice D 2905.

6. Sources of Data

6.1 Plan and conduct an interlaboratory study as directed in Practice D 2904 or in Practice E 691 to secure the information needed to estimate the necessary components of variance. For new test methods for which an interlaboratory test has not been run, see Section 16.

7. Categories of Data

7.1 *General*—Individual observations obtained as directed in a test method fall into a number of categories that require different treatments of the data. The more important of such categories are discussed in standard statistical texts **(1), (2), (3), (4), (5), (6),** and **(7)** and are briefly described in the following sections:

7.2 *Normal Distribution*—If the frequency distribution of individual observations approximates the normal curve and the size of the standard deviation is independent of the average level of the observations, the data can be assumed to be normally distributed and the standard deviation should be used as the measure of variability. Generally, frequency distributions having a hump somewhere near the middle of the distribution and tailing off on either side of the hump approach the normal curve closely enough to permit using data handling techniques based on the normal curve without seriously distorting the conclusions.

NOTE 5—It is recommended that qualified assistance be sought when data do not conform to the normal distribution, when the response is not the arithmetic average of the observations, or both. Within ASTM Committee D-13 such help is available through Subcommittee D13.93 on Statistics.

7.3 *Transformed Data*—If the individual observations have a frequency distribution that is markedly skewed, if the standard deviation seems to be correlated with the average of the observations, or if both these conditions exist; consider transforming the original data to obtain values that are approximately normally distributed with a standard deviation that is independent of the average. Arbitrary grades or classifications and scores of ranked data are among the types of data that usually require transformation before they can be treated as being normally distributed variables.

NOTE 6—In the case of arbitrary grades or classifications and of scores for ranked data, the observations may have such a complex nonlinear relationship that meaningful transformations may not be practicable. If this is so, precision statements must be based on subjective judgement rather than on the analysis of observed data.

NOTE 7—An empirically chosen transformation is often considered satisfactory if a cumulative frequency distribution of the transformed data gives a reasonably straight line when plotted on normal probability graph paper.[8] A number of articles and standard statistical texts discuss the choice of suitable transformations **(2), (3), (8), (9),** and **(10).** (See also Practice D 2904.)

7.3.1 When the shape of the distribution of individual observations is reasonably symmetrical but the standard deviation is proportional to the average of the observations, consider converting the standard deviation to the coefficient of variation, *CV* %, using Eq 1 :

$$CV\,\% = 100\,s / \bar{X} \qquad (1)$$

where:
CV % = coefficient of variation as a percent of the average,
s = standard deviation in units of measure, and

[8] Normal probability graph paper may be bought from most suppliers. The equivalent of Keuffel and Esser Co. Style 46-8000 or of Codex Book Co., Inc., Norwood, MA 02062, Style 3127, is acceptable.

$\bar{X}$ = average of all the observations for a specific material.

NOTE 8—Because the transformation is made on the standard deviation and not on the individual observations, the coefficient of variation is not always recognized as a transformation. The same results can be obtained, however, by transforming the individual observations.

NOTE 9—Use of the coefficient of variation when the standard deviation is the more appropriate measure of variability can cause serious errors. Likewise, the use of the standard deviation when the coefficient of variation is the more appropriate measure can result in serious errors.

7.4 *Binomial Distribution*—The binomial distribution applies to test results that are discrete variables reporting the number of successes or failures in a specified number of observations. Each observation in such a test result is an attribute; that is, a nonnumerical report of success or failure based on criteria specified in the procedure (see 7.6).

7.5 *Poisson Distribution*—The Poisson distribution applies to test results that report a count of the number of incidents, such as a specified type of defect, observed over a specified period of time or in one or more specimens of a specified size. The observed count in a Poisson distribution must be small in comparison to the potential count. Examples of data having Poisson distributions are the number of defects of a specified type counted in a specified area of a textile material and the number of stops or other incidents reported for a specified block of equipment over a specified time span.

7.6 *Attributes*—No justifiable statement can be made about the precision or the bias of a test result that is an attribute; that is, a nonnumerical report of success or failure based on criteria specified in the procedure. Test results that are a number summarizing the attributes of individual observations are discrete variables about which justifiable statements can be made on precision and bias (see 7.4 and 7.5).

8. Calculations for Normal Distributions and Transformed Data

8.1 *General*—The same calculations are required for normal distributions having variability measured by standard deviations and for all distributions for which the data have been transformed in order to approach a normal distribution, to make the measure of variability independent of the average, or to obtain both of these objectives. The use of the coefficient of variation as a substitute for the standard deviation is also covered.

8.2 *Calculating Critical Differences*— Calculate the critical differences for averages of observations using Eq 2 or Eq 3:

$$\text{Critical difference between averages, units of measure} = 1.414\, z\, s_T \quad (2)$$

$$\text{Critical difference between averages, percent of average} = 1.414\, z\, v_T \quad (3)$$

where:

1.414 = square root of 2, a constant that converts the standard error of an average to the standard error for the difference between two such averages,

z = standard normal deviate for two-sided limits and the specified probability level (z = 1.960 for the 95 % probability level),

s_T = standard error for the specific size and type of averages being compared (see 8.4), and

v_T = coefficient of variation for the specific size and type of averages being compared (see 8.4).

NOTE 10—Generally, infinite degrees of freedom are assumed when calculating critical differences and confidence limits based on the best information obtainable from existing interlaboratory tests. There are reasonable statistical arguments for this usage. Even if the degrees of freedom associated with each component of variance has been calculated by Satterthwaite's approximation **(1)** or a comparable procedure, there are no generally accepted methods known for assigning degrees of freedom to a standard error which combines two or more components of variance, each having a different number of degrees of freedom, as is done in Eq 7 and 8 and in Eq 10 and 11.

8.3 *Calculating Confidence Limits*— Calculate the width of the confidence limits for averages of observations using Eq 4 or Eq 5:

$$\text{Width of confidence limits for averages, units of measure} = \pm z\, s_T \quad (4)$$

$$\text{Width of confidence limits for averages, percent of average} = \pm z\, v_T \quad (5)$$

where the terms in the equations are defined in 8.2.

NOTE 11—The property being evaluated may need to be controlled in only one direction; for example, excessive fabric shrinkage is important but too little shrinkage is not likely to be undesirable. Nevertheless, "plus and minus" confidence limits are suggested even in these cases since confidence limits are normally used to express the variability in a single average. Critical differences should be used to compare pairs of averages.

8.4 *Combining Components of Variance*— Calculate the standard error of the specific size and type of averages that are to be compared using Eq 6, Eq 7, or Eq 8:

$$s_T\ (\text{single}-\text{operator}) = (s_s^2/n)^{1/2} \quad (6)$$

$$s_T\ (\text{within}-\text{laboratory}) = [s_w^2 + (s_s^2/n)]^{1/2} \quad (7)$$

$$s_T\ (\text{between}-\text{laboratory}) = [s_B^2 + s_w^2 + (s_s^2/n)]^{1/2} \quad (8)$$

where the equations respectively calculate the standard error of averages of observations under the conditions of single-operator, within-laboratory, and between-laboratory precision, and

where:

s_S^2 = single-operator component of variance or the residual error component of variance,

s_W^2 = within-laboratory component of variance (Note 12),

s_B^2 = between-laboratory component of variance, and

n = number of observations by a single operator in each average.

NOTE 12—The within-laboratory component of variance is the sum of all the individual components of variance, except the single-operator component of variance, that contribute to the variability of observations within a single laboratory. Included are such components of variance as those for days of testing, units of apparatus, and different operators within a single laboratory. If the within-laboratory component of variance is not calculated separately, all sources of variability except the single-operator component are included in the between-laboratory component. Under these conditions, calculate the standard error (between-laboratory) using zero for the within-laboratory component.

When an interlaboratory test program run as directed in Practice D 2904 results in a significant material by laboratory interaction or a material by

operator (within laboratories) interaction, see Annex A1.14.2 of Practice D 2904 for instructions on estimating the components of variance for multi-material comparisons.

If components of variance are to be expressed as coefficients of variation, calculate them using Eq 1 and calculate the coefficient of variation for the specific size and type of averages that are to be compared using Eq 9, Eq 10, or Eq 11:

$$v_T \text{ (single−operator)} = (v_s^2/n)^{1/2} \quad (9)$$

$$v_T \text{ (within−laboratory)} = [v_w^2 + (v_s^2/n)]^{1/2} \quad (10)$$

$$v_T \text{ (between laboratory)} = [v_B^2 + v_w^2 + (v_s^2/n)]^{1/2} \quad (11)$$

where the meanings of the subscripts for the individual components of variance expressed as coefficients of variation are the same as in the legend for Eq 6, Eq 7, and Eq 8.

8.5 *Sample Calculations—Components of Variance as Standard Deviations*:

8.5.1 *Example 1: Single-Operator Precision*—At the 95% probability level, calculate the critical difference and confidence limits for averages of ten observations when the single-operator component of variance expressed as a standard deviation is 1.8 percentage points. Using Eq 6, $s_T = [(1.8)^2/10]^{1/2} = 0.57$ percentage points. Using Eq 2, the critical difference = $1.414 \times 1.960 \times 0.57 = 1.58$ percentage points. Using Eq 4, the width of the confidence limits = $\pm(1.960 \times 0.57) = \pm 1.12$ percentage points (Note 12).

8.5.2 *Example 2: Within-Laboratory Precision (Multi-Operator)*—At the 95 % probability level, calculate the critical difference and confidence limits for averages of ten when the single-operator and within-laboratory components of variance expressed as standard deviations are respectively 1.8 and 0.3 percentage points. Using Eq 7, $s_T = [(0.3)^2 + ((1.8)^2/10)]^{1/2} = 0.64$ percentage points. Using Eq 2, the critical difference = $1.414 \times 1.960 \times 0.64 = 1.77$ percentage points. Using Eq 4, the width of the confidence limits = $\pm(1.960 \times 0.64) = \pm 1.25$ percentage points (Note 12).

8.5.3 *Example 3: Between-Laboratory Precision*—At the 95 % probability level, calculate the critical difference and the confidence limits for averages of ten when the single-operator, within-laboratory, and between-laboratory components of variance expressed as standard deviations are respectively 1.8, 0.3, and 0.5 percentage points. Using Eq 8, $s_T = [(0.5)^2 + (0.3)^2 + ((1.8)^2/10)]^{1/2} = 0.81$ percentage points. Using Eq 2, the critical difference = $1.414 \times 1.960 \times 0.81 = 2.24$ percentage points. Using Eq 4, the width of the confidence limits = $\pm(1.960 \times 0.81) = \pm 1.59$ percentage points (Note 12).

TABLE 1 Components of Variance as Standard Deviations, Percentage Points

Names of the Properties	Single-Operator Component	Within-Laboratory Component	Between-Laboratory Component
(Insert here name of Property 1)	1.8	0.3	0.5
(Insert here name of Property 2)	1.2	0.4	0.0

8.6 *Sample Calculations—Components of Variance as Coefficients of Variation*:

8.6.1 *Example 4: Within-Laboratory Precision*—At the 95 % probability level, calculate the critical difference and the confidence limits for averages of five observations when the single-operator component of variance expressed as a coefficient of variation is 5.3 % of the average. Using Eq 9, $v_T = [(5.3)^2/5]^{1/2} = 2.37$ % of the average. Using Eq 3, the critical difference = $1.414 \times 1.960 \times 2.37 = 6.57$ % of the average. Using Eq 5, the width of the confidence limits = $\pm(1.960 \times 2.37) = \pm 4.65$ % of the average. (Note 12).

8.6.2 *Example 5: Within-Laboratory Precision (Multi-Operator)*—At the 95 % probability level, calculate the critical difference and the confidence limits for averages of five observations when the single-operator and within-laboratory components of variance expressed as coefficients of variation are respectively 5.3 and 1.0 % of the average. Using Eq 10, $v_T = [(1.0)^2 + ((5.3)^2/5)]^{1/2} = 2.57$ % of the average. Using Eq 3, the critical difference = $1.414 \times 1.960 \times 2.57 = 7.12$ % of the average. Using Eq 5, the width of the confidence limits = $\pm(1.960 \times 2.57) = \pm 5.04$ % of the average (Note 12).

8.6.3 *Example 6: Between-Laboratory Precision*—At the 95 % probability level, calculate the critical difference and the confidence limits for averages of five observations when the single-operator, within-laboratory, and between-laboratory components of variance are expressed as coefficients of variation and are respectively 5.3, 1.0, and 2.0 % of the average. Using Eq 11, $v_T = [(2.0)^2 + (1.0)^2 + ((5.3)^2/5)]^{1/2} = 3.26$ % of the average. Using Eq 3, the critical difference = $1.414 \times 1.960 \times 3.26 = 9.03$ % of the average. Using Eq 5, the width of the confidence limits = $\pm(1.960 \times 3.26) = \pm 6.39$ % of the average.

9. Calculations for Binomial Distributions

9.1 *Critical Differences for Binomial Distributions*—Determine critical differences between two test results from a binomial distribution using an exact test of significance for 2 by 2 contingency tables containing small frequencies. Prepare a table of critical differences using published tables (Table 28, Ref **7**), the methods shown in 9.1.1 and 9.1.2, or an algorithm for use with a computer.[9] See Table 3 for an example of such a table.

NOTE 13—For data from both the binomial and Poisson distributions, the tables of critical differences and of confidence limits are based on the mathematical characteristics of the applicable frequency distribution. Bias in actual test results due to systematic errors in testing for some or all of the observations will normally have the effect of reducing the actual probability level to some unknown value which is less than the value shown in the tables.

9.1.1 Calculate the probability of reporting exactly *a* successes in one of the test results using Eq 12:

[9] Printouts and punched cards describing all of the algorithms mentioned in this recommended practice are available from ASTM Headquarters, 100 Barr Harbor Drive, West Conshohocken, PA 19428, at a nominal cost. Request ADJD2906.