



Standard Practice for Calculating and Using Basic Statistics¹

This standard is issued under the fixed designation E2586; 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.

1. Scope

1.1 This practice covers methods and equations for computing and presenting basic descriptive statistics using a set of sample data containing a single variable. This practice includes simple descriptive statistics for variable data, tabular and graphical methods for variable data, and methods for summarizing simple attribute data. Some interpretation and guidance for use is also included.

1.2 The system of units for this practice is not specified. Dimensional quantities in the practice are presented only as illustrations of calculation methods. The examples are not binding on products or test methods treated.

1.3 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 ASTM Standards:²

E178 Practice for Dealing With Outlying Observations

E456 Terminology Relating to Quality and Statistics

E2282 Guide for Defining the Test Result of a Test Method

E3080 Practice for Regression Analysis

2.2 ISO Standards:³

ISO 3534-1 Statistics—Vocabulary and Symbols, part 1: Probability and General Statistical Terms

ISO 3534-2 Statistics—Vocabulary and Symbols, part 2: Applied Statistics

3. Terminology

3.1 *Definitions*—Unless otherwise noted, terms relating to quality and statistics are as defined in Terminology E456.

¹ This practice is under the jurisdiction of ASTM Committee E11 on Quality and Statistics and is the direct responsibility of Subcommittee E11.10 on Sampling / Statistics.

Current edition approved Nov. 1, 2016. Published November 2016. Originally approved in 2007. Last previous edition approved in 2014 as E2586 – 14. DOI: 10.1520/E2586-16.

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>.

3.1.1 *characteristic, n*—a property of items in a sample or population which, when measured, counted, or otherwise observed, helps to distinguish among the items. **E2282**

3.1.2 *coefficient of variation, CV, n*—for a nonnegative characteristic, the ratio of the standard deviation to the mean for a population or sample

3.1.2.1 *Discussion*—The coefficient of variation is often expressed as a percentage.

3.1.2.2 *Discussion*—This statistic is also known as the *relative standard deviation, RSD*.

3.1.3 *confidence bound, n*—see *confidence limit*.

3.1.4 *confidence coefficient, n*—see *confidence level*.

3.1.5 *confidence interval, n*—an interval estimate [L, U] with the statistics L and U as limits for the parameter θ and with confidence level $1 - \alpha$, where $\Pr(L \leq \theta \leq U) \geq 1 - \alpha$.

3.1.5.1 *Discussion*—The confidence level, $1 - \alpha$, reflects the proportion of cases that the confidence interval [L, U] would contain or cover the true parameter value in a series of repeated random samples under identical conditions. Once L and U are given values, the resulting confidence interval either does or does not contain it. In this sense “confidence” applies not to the particular interval but only to the long run proportion of cases when repeating the procedure many times.

3.1.6 *confidence level, n*—the value, $1 - \alpha$, of the probability associated with a confidence interval, often expressed as a percentage.

3.1.6.1 *Discussion*— α is generally a small number. Confidence level is often 95 % or 99 %.

3.1.7 *confidence limit, n*—each of the limits, L and U, of a confidence interval, or the limit of a one-sided confidence interval.

3.1.8 *degrees of freedom, n*—the number of independent data points minus the number of parameters that have to be estimated before calculating the variance.

3.1.9 *estimate, n*—sample statistic used to approximate a population parameter.

3.1.10 *histogram, n*—graphical representation of the frequency distribution of a characteristic consisting of a set of rectangles with area proportional to the frequency. **ISO 3534-1**

3.1.10.1 *Discussion*—While not required, equal bar or class widths are recommended for histograms.