Designation: E 1970 – 01

Standard Practice for Statistical Treatment of Thermoanalytical Data[1]

This standard is issued under the fixed designation E 1970; 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 details the statistical data treatment used in some thermal analysis methods.

1.2 The method describes the commonly encountered statistical tools of the mean, standard derivation, relative standard deviation, pooled standard deviation, pooled relative standard deviation and the best fit to a straight line, all calculations encountered in thermal analysis methods.

1.3 Some thermal analysis methods derive the analytical value from the slope or intercept of a best fit straight line assigned to three or more sets of data pairs. Such methods may require an estimation of the precision in the determined slope or intercept. The determination of this precision is not a common statistical tool. This practice details the process for obtaining such information about precision.

1.4 Computer or electronic-based instruments, techniques or data treatment equivalent to this practice may also be used.

NOTE 1—Users of this practice are expressly advised that some such instruments or techniques may not be equivalent. It is the responsibility of the user of this standard to determine the necessary equivalency prior to use.

1.5 SI units are the standard.

1.6 There are no ISO methods equivalent to this practice.

2. Referenced Documents

2.1 *ASTM Standards:*

E 177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods[2]

E 456 Terminology Relating to Quality and Statistics[2]

3. Terminology

3.1 *Definitions*—The technical terms used in this practice are defined in Practice E 177 and Terminology E 456.

3.2 *Symbols:*[3]

m = slope
b = intercept
n = number of data sets (that is, x_i, y_i)
x_i = an individual independent variable observation
y_i = an individual dependent variable observation
Σ = mathematical operation which means "the sum of all" for the term(s) following the operator
X = mean value
s = standard deviation
s_{pooled} = pooled standard deviation
s_b = standard deviation of the line intercept
s_m = standard deviation of the slope of a line
s_y = standard deviation of Y values
RSD = relative standard deviation
δy_i = variance in y parameter
r = correlation coefficient

4. Summary of Practice

4.1 The result of a series of replicate measurements of a value are typically reported as the mean value plus some estimation of the precision in the mean value. The standard deviation is the most commonly encountered tool for estimating precision, but other tools, such as relative standard deviation or pooled standard deviation, also may be encountered in specific thermoanalytical test methods. This practice describes the mathematical process of achieving mean value, standard deviation, relative standard deviation and pooled standard deviation.

4.2 In some thermal analysis experiments, a linear or a straight line, response is assumed and desired values are obtained from the slope or intercept of the straight line through the experimental data. In any practical experiment, however, there will be some uncertainty in the data so that results are scattered about such a straight line. The least squares method is an objective tool for determining the "best fit" straight line drawn through a set of experimental results and for obtaining information concerning the precision of determined values.

4.2.1 For the purposes of this practice, it is assumed that the physical behavior, which the experimental results approximate, are linear with respect to the controlled value, and may be represented by the algebraic function:

$$y = mx + b \quad (1)$$

4.2.2 Experimental results are gathered in pairs, that is, for

[1] This practice is under the jurisdiction of ASTM Committee E37 on Thermal Measurements and is the direct responsibility of Subcommittee E37.01 on Test Methods and Recommended Practices.

Current edition approved Aug. 10, 2001. Published November 2001. Originally published as E 1970 – 98. Last previous edition E 1970 – 00.

[2] *Annual Book of ASTM Standards*, Vol 14.02.

[3] Taylor, J.K., *Handbook for SRM Users*, Publication 260-100, National Institute of Standards and Technology, Gaithersburg, MD, 1993.

every corresponding x_i (controlled) value, there is a corresponding y_i (response) value.

4.2.3 The best fit approach assumes that all x_i values are exact and the y_i values (only) are subject to uncertainty.

NOTE 2—In experimental practice, both x and y values are subject to uncertainty. If the uncertainty in x_i and y_i are of the same relative order of magnitude, other more elaborate fitting methods should be considered. For many sets of data, however, the results obtained by use of the assumption of exact values for the x_i data constitute such a close approximation to those obtained by the more elaborate methods that the extra work and additional complexity of the latter is hardly justified.[4]

4.2.4 The best fit approach seeks a straight line, which minimizes the uncertainty in the y_i value.

5. Significance and Use

5.1 The standard deviation, or one of its derivatives, such as relative standard deviation or pooled standard deviation, derived from this practice, provides an estimate of precision in a measured value. Such results are ordinarily expressed as the mean value ± the standard deviation, that is, $X \pm s$.

5.2 If the measured values are, in the statistical sense, "normally" distributed about their mean, then the meaning of the standard deviation is that there is a 67 % chance, that is 2 in 3, that a given value will lie within the range of ± one standard deviation of the mean value. Similarly, there is a 95 % chance, that is 19 in 20, that a given value will lie within the range of ± two standard deviations of the mean. The two standard deviation range is sometimes used as a test for outlying measurements.

5.3 The calculation of precision in the slope and intercept of a line, derived from experimental data, commonly is required in the determination of kinetic parameters, vapor pressure or enthalpy of vaporization. This practice describes how to obtain these and other statistically derived values associated with measurements by thermal analysis.

6. Calculation

6.1 Commonly encountered statistical results in thermal analysis are obtained in the following manner.

NOTE 3—In the calculation of intermediate or final results, all available figures shall be retained with any rounding to take place only at the expression of the final results according to specific instructions or to be consistent with the precision and bias statement.

6.1.1 The mean value (X) is given by:

$$X = \frac{x_1 + x_2 + x_3 + \ldots + x_i}{n} = \frac{\Sigma x_i}{n} \tag{2}$$

6.1.2 The standard deviation (s) is given by:

$$s = \left[\frac{\Sigma(x_i - X)^2}{(n-1)}\right]^{1/2} \tag{3}$$

6.1.3 The Relative Standard Deviation (RSD) is given by:

$$RSD = (s \cdot 100\ \%) / X \tag{4}$$

6.1.4 The Pooled Standard Deviation (s_p) is given by:

$$s_p = \left[\frac{(\{n_1 - 1\} \cdot s_1^2) + (\{n_2 - 1\} \cdot s_2^2) + \ldots + \Sigma(\{n_i - 1\} \cdot s_i^2)}{(n_1 - 1) + (n_2 - 1) + \ldots + (n_i - 1)}\right]^{1/2} \tag{5}$$

$$= \left[\frac{\Sigma(\{n_i - 1\} \cdot s_i)}{\Sigma(n_i - 1)}\right]^{1/2} \tag{6}$$

NOTE 4—For the calculation of pooled relative standard deviation, the values of s_i are replaced by RSD_i.

6.2 *Best Fit to a Straight Line*:

6.2.1 The best fit slope (m) is given by:

$$m = \frac{n\Sigma(x_i y_i) - (\Sigma x_i)(\Sigma y_i)}{n\Sigma x_i^2 - (\Sigma x_i)^2} \tag{7}$$

6.2.2 The best fit intercept (b) is given by:

$$b = \frac{(\Sigma x_i^2)(\Sigma y_i) - (\Sigma x_i)(\Sigma x_i y_i)}{n\Sigma x_i^2 - (\Sigma x_i)^2} \tag{8}$$

6.2.3 The individual dependent parameter variance (δy_i) of the dependent variable (y_i) is given by:

$$\delta y_i = y_i - (mx_i + b) \tag{9}$$

6.2.4 The standard deviation s_y of the set of y values is given by:

$$s_y = \left[\frac{\Sigma(\delta y_i)^2}{n-2}\right]^{1/2} \tag{10}$$

6.2.5 The standard deviation (s_m) of the slope is given by:

$$s_m = s_y\left[\frac{n}{n\Sigma x_i^2 - (\Sigma x_i)^2}\right]^{1/2} \tag{11}$$

6.2.6 The standard deviation (s_b) of the intercept (b) is given by:

$$s_b = s_y\left[\frac{\Sigma x_i^2}{n\Sigma x_i^2 - (\Sigma x_i)^2}\right]^{1/2} \tag{12}$$

6.2.7 The denominators in Eqs 6, 7, 9, and 10 are the same. It is convenient to obtain the denominator (D) as a separate function for use in manual calculation of each of these equations.

$$D = n\Sigma x_i^2 - (\Sigma x_i)^2 \tag{13}$$

6.2.8 The linear correlation coefficient (r), a measure of the mutual dependence between paired x and y values, is given by:

$$r = \frac{n\Sigma xy - (\Sigma x_i)(\Sigma y_i)}{\{[n\Sigma x_i^2 - (\Sigma x_i)^2]^{1/2}\ [n(\Sigma y_i^2) - (\Sigma y_i)^2]^{1/2}\}} \tag{14}$$

NOTE 5—r may vary from –1 to + 1, where values of + or –1 indicate perfect (100 %) correlation and 0 indicates no (0 %) correlation, that is, random scatter. A positive (+) value indicates a positive slope and a negative (-) indicates a negative slope.

6.3 *Example Calculations*:

6.3.1 Table 1 provides an example set of data and intermediate calculations which may be used to examine the manual calculation of slope (m) and its standard deviation (s_m) and of the intercept (b) and its standard deviation (s_b).

6.3.1.1 The values in Columns A and B are experimental parameters with x_i being the independent parameter and y_i the dependent parameter.

6.3.1.2 From the individual values of x_i and y_i in Columns A and B in Table 1, the values for x_i^2 and $x_i y_i$ are calculated and placed in Columns C and D.

[4] Mandel, J., *The Statistical Analysis of Experimental Data*, Dover Publications, New York, NY, 1964.

TABLE 1 Example Set of Data and Intermediate Calculations

Column Experi-ment	A x_i	B y_i	C x_i^2	D $x_i\,y_i$	E m x_i + b	F δy_i	G $(\delta y_i)^2$	H $(y_i)^2$
1	1.0	1.2	1.0	1.2	1.1997	0.0003	0.000 000 09	1.44
2	1.0	1.3	1.0	1.3	1.1997	0.1003	0.010 060 09	1.69
3	12.0	13.7	144.0	164.0	13.6924	0.0076	0.000 057 76	187.69
4	12.0	13.5	144.0	162.0	13.6924	-0.1924	0.037 017 76	182.25
5	25.0	28.5	625.0	712.5	28.4565	0.0435	0.001 892 25	812.25
6	25.0	28.5	625.0	712.5	28.4565	0.0435	0.001 892 25	812.25
Σ	76.0	86.7	1540.0	1753.9			0.050 920 20	1997.57

NOTE 1—n = 6.

6.3.1.3 The values in columns A, B, C, and D are summed (added) to obtain $\Sigma x_i = 76.0$, $\Sigma_{yi} = 86.7$, $\Sigma x_i^2 = 1540.0$, and $\Sigma x_i y_i = 1753.9$, respectively.

6.3.1.4 The denominator (D) is calculated using Eq 13 and the values $\Sigma x_i^2 = 1540.0$ and $\Sigma x_i = 76.0$ from 6.3.1.3.

$$D = (6 \cdot 1540.0) - (76.0 \cdot 76.0) = 3464.0 \tag{15}$$

6.3.1.5 The value for m is calculated using the values $n = 6$ $\Sigma x_i \cdot y_i = 1753.9$, $\Sigma x_i = 76.0$, $\Sigma y_i = 86.7$, and $D = 3640.0$, from 6.3.1.3 and 6.3.1.4 and Eq 6:

$$m = \frac{n\Sigma(x_i\,y_i) - \Sigma x_i\,\Sigma y_i}{D} \tag{16}$$

$$\begin{aligned} m &= \frac{(6 \cdot 1753.9) - (76.0 \cdot 86.7)}{3464.0} \\ &= \frac{10523.4 - 6589.2}{3464.0} \\ &= 1.1357 \end{aligned} \tag{17}$$

6.3.1.6 The value for b is calculated using the values $n = 6$, $\Sigma x_i \cdot y_i = 1753.9$, $\Sigma x_i = 76.0$, and $\Sigma_{yi} = 86.7$, from 6.3.1.3 and 6.3.1.4 and Eq 7:

$$\begin{aligned} b &= \frac{(1540.0 \cdot 86.7) - (76.0 \cdot 1753.9)}{3464.0} \\ &= \frac{133518.0 - 133296.4}{3464.0} \\ &= 0.064 \end{aligned} \tag{18}$$

6.3.1.7 Using the values for $m = 1.1357$ and $b = 0.064$ from 6.3.1.5 and 6.3.1.6, and the value $\Sigma x_i = 76.0$ from Table 1, the $n = 6$, values for δy_i are calculated values using Eq 8 and recorded in Column F in Table 1.

6.3.1.8 From the values in Column F of Table 1, the six values for $(\delta y_i)^2$ are calculated and recorded in Column G.

6.3.1.9 The values in Column G of Table 1 are summed to obtain $\Sigma\ (\delta y_i)^2$.

6.3.1.10 The value of s_y is calculated using the value from 6.3.1.9 and Eq 10:

$$s_y = [0.050\,092\,02 / 4]^{1/2} = 0.1119 \tag{19}$$

6.3.1.11 The value for s_m (expressed to two significant figures) is calculated using the values of $D = 3464.0$ and $s_y = 0.1119$ from 6.3.1.4 and 6.3.1.10, respectively.

$$s_m = 0.1119 \left[\frac{6}{3464.0}\right]^{1/2} = 0.0047 \tag{20}$$

6.3.1.12 The value for s_b (expressed to two significant figures) is calculated using the values ofΣx_i^2, $D = 3464.0$, and $s_y = 0.119$, from 6.3.1.3, 6.3.1.4, and 6.3.1.10, respectively.

$$s_b = 0.1119 \left[\frac{1540.0}{3464.0}\right]^{1/2} = 0.075 \tag{21}$$

6.3.1.13 The value of the slope along with its estimation of precision is obtained from 6.3.1.5 and 6.3.1.11 and reported as follows:

$$m \pm s_m \tag{22}$$

$$m = 1.1357 \pm 0.0047 \tag{23}$$

6.3.2 Table 1 provides an example set of data that may be used to examine the manual calculation of the correlation coefficient (r).

6.3.2.1 The value of r is calculated using the values n = 6, $\Sigma x_i = 76.0$, $\Sigma y_i = 86.7$, $\Sigma x_i^2 = 1540.0$, $\Sigma x_i y_i = 1753.9$, and $\Sigma (y_i)^2 = 1997.57$ from Table 1 and Eq 13.

$$r = \frac{\{(6 \cdot 1753.9) - (76.0 \cdot 86.7)\}}{\{[(6 \cdot 1540.0) - (76.0 \cdot 76.0)]^{1/2} \cdot [(6 \cdot 1997.57) - (86.7 \cdot 86.7)]^{1/2}\}} \tag{24}$$

$$\begin{aligned} &= \frac{\{10523.4 - 6589.2\}}{\{[9240 - 5776]^{1/2} \cdot [11985.42 - 7516.89]^{1/2}\}} \\ &= \frac{3934.2}{\{[3464]^{1/2} \cdot [4468.53]^{1/2}\}} \\ &= \frac{3934.2}{\{58.856 \cdot 66.847\}} \\ &= 0.99996 \end{aligned}$$

7. Report

7.1 Report the following information:

7.1.1 All of the statistical values required to meet the needs of the respective applications method.

7.1.2 The specific dated version of this practice that is used.

8. Keywords

8.1 intercept; mean; precision; relative standard deviation; slope; standard deviation