



Designation: E 876 – 89 (Reapproved 1994)^{ε1}

Standard Practice for Use of Statistics in the Evaluation of Spectrometric Data¹

This standard is issued under the fixed designation E 876; 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.

ε¹ NOTE—Section 7 was added editorially in January 1995.

1. Scope

1.1 This practice provides for the statistical evaluation of data obtained from spectrometrical methods of analysis. Included are definitions used in statistics, methods to determine variance and standard deviation of data, and calculations for (1) estimate of variance and pooling estimates of variance, (2) standard deviation and relative standard deviation, (3) testing for outliers, (4) testing for bias, (5) establishing limits of detection, and (6) testing for drift.

2. Referenced Documents

2.1 ASTM Standards:

E 135 Terminology Relating to Analytical Chemistry for Metals, Ores, and Related Materials²

E 177 Practice for Use of the Terms Precision and Bias in ASTM Test Methods³

E 178 Practice for Dealing with Outlying Observations³

E 305 Practice for Establishing and Controlling Spectrochemical Analytical Curves²

E 456 Terminology Relating to Quality and Statistics³

3. Terminology

3.1 Definitions:

3.1.1 For definitions of terms used in this practice, refer to Terminologies E 135 and E 456.

3.1.2 All quantities computed from limited data are defined as estimates of the parameters that are properties of the system (population) from which the data were obtained.

3.2 Definitions of Terms Specific to This Standard:

3.2.1 *average measurement* ($\bar{x}$)—the arithmetic mean obtained by dividing the sum of the measurements by the number of measurements. It is an estimate of μ , the value of the population that the average would become if the number of measurements were infinite. Either $\bar{x}$ or μ may include a systematic error if there is a bias in the measurement.

3.2.2 *bias*—a systematic displacement of all or most determinations from the assumed true value. An acceptable bias should be agreed upon prior to testing a method. *Accuracy*, often used to qualify a method, is a measurement which includes both imprecision and bias.

NOTE 1—*Precision* and *bias* are discussed in detail in Practice E 177. In analytical methods, *precision* refers to the distribution of repeat determinations about the average. All analyses are presumed to have been made under the same set of conditions. Standard deviation provides a measure of this distribution.

NOTE 2—An evaluation of a method will be sample-dependent. Multiple samples should be tested for homogeneity since even certified reference materials may exhibit significantly different degrees of inhomogeneity. A measure of both sample and method precision may be made by replicating determinations on specific portions of the sample specimens.

3.2.3 *confidence to be placed on the estimate of μ* (μ)—the average, $\bar{x}$, is expected to be close to μ and should be very close if the number of determinations is large, no significant bias exists and the standard deviation, s , is small. The degree of closeness is expressed as a probability (*confidence level*) that μ is in a specified interval (*confidence interval*) centered at $\bar{x}$. With a certain probability, limits are placed on the quantity $\bar{x}$ which may include the unknown quantity μ . A probability level, p %, can be selected so that μ will be within the limits placed about $\bar{x}$. See 3.2.1

3.2.4 *degrees of freedom* (df)—the number of contributors to the deviations of a measurement. Since a deviation can be implied only when there are at least two members of a group, the degrees of freedom of a set of measurements is generally one less than the number of measurements. It is the sample size less the number of parameters estimated. If the group is a listing of a series of differences of measurements or a series of determinations of variance, the degrees of freedom is the number of these differences or the total of the degrees of freedom of each series of determinations.

3.2.5 *detection limit*—paraphrasing the definition in Terminology E 135, it is the lowest estimated concentration that permits a confident decision that an element is present. The actual concentration being measured falls within a confidence interval that encompasses the estimated concentration. The lowest estimate has a confidence interval that reaches to zero concentration, but not below. It cannot be assumed that the estimated concentration is an actual concentration. Neither can

¹ This practice is under the jurisdiction of ASTM Committee E-1 on Analytical Chemistry for Metals, Ores and Related Materials and is the direct responsibility of Subcommittee E01.22 on Statistics and Quality Control.

Current edition approved Nov. 20, 1989. Published January 1990. Originally published as E 876 – 82. Last previous edition E 876 – 89.

² *Annual Book of ASTM Standards*, Vol 03.05.

³ *Annual Book of ASTM Standards*, Vol 14.02.



it be assumed that an actual concentration that equals the detection limit will always give a positive detection. The definition in Terminology E 135 properly characterizes this detection as being a limiting value. See also *determination limit* which follows.

3.2.6 *determination limit*—the estimated low concentration where the range of the encompassing confidence interval bears some specified maximum ratio to that concentration. The ratio would depend upon what is acceptable in a specific application.

3.2.7 *drift*—a gradual, systematic change in measurements (either increasing or decreasing) from start to completion of a set of replicate determinations of the same material.

3.2.8 *estimate of standard deviation (s)*—the square root of the estimate of variance. It is a measure of the variability of a set of values representing the whole population. It is an estimate of σ , the actual standard deviation of an infinite number of measurements. With normal distribution, 68 % of the values in a population will fall within $\pm\sigma$ of the true value, μ ; 95 % within $\pm 2 \sigma$ of μ ; and 99.7 % within $\pm 3 \sigma$ of μ .

3.2.9 *estimate of variance (s²)*—a measure of precision of a measurement based on summing the squares of the deviations of individual determinations from the average and dividing by the degrees of freedom.

3.2.10 *outlier*—a measurement that, for a specific degree of confidence, is not part of the population.

3.2.11 *pooled estimate of variance (s_p²)*—the combined estimate of variance calculated from two or more estimates under the same or similar conditions. Pooling estimates increases the degrees of freedom and improves the quality of the estimate if the variance is approximately the same for each measurement.

NOTE 3—If the concentration level varies considerably within the pooled data, the pooled variance may be inaccurate. It may be possible in such cases, however, to determine a valid estimate by pooling relative standard deviations.

3.2.12 *precision*—the agreement among repeat measurements, usually expressed as either repeatability or reproducibility as defined in Terminology E 456 (see Note 1 and Note 2).

3.2.13 *range (w)*—the difference between the highest and lowest measurements for a series of values obtained under identical conditions. Range is useful for estimating standard deviation and for determining if certain values are outliers.

3.2.14 *relative standard deviation (RSD)*—the standard deviation as a percentage of the average analysis or reading. By providing a means of expressing the precision in relative rather than absolute terms, it may serve to show a more consistent measure of precision for widely different values of $\bar{x}$.

3.2.15 *standard error*—a term sometimes used synonymously with standard deviation but which will be used here to measure how consistently the accepted true concentrations of a series of reference materials compare to the apparent concentrations determined from a calibration. It is an estimate that is similar to standard deviation except for the degrees of freedom used. When the measurement is used to define the effectiveness of the calibration established by these reference materials, the degrees of freedom is the number of data points minus the number of constants in the calibration (the sample size less the

number of parameters estimated). If correction is made for spectral interference, the number of constants used in the correction should be counted as being calibration constants. This would be true whether correction was made to readings or to final concentrations. When the measurement is made with new data and applied to a previously determined calibration, the degrees of freedom is the number of references used to make the test. In either case, the significance of the measurement is limited to the range of concentration in the reference materials and implies that concentrations will be fairly well spread within that range.

4. Significance and Use

4.1 The data obtained in spectrometrical analyses may be evaluated as statistical measurements. Use of the various determinations of precision which follow permits a consistent basis for comparing methods of analysis or for monitoring their performance.

4.2 Some explanations are included to clarify the function of the statistical calculations being made.

4.3 Examples of all calculations are given in the appendixes.

5. Calculation

5.1 Average ($\bar{x}$)

$$\bar{x} = \Sigma x/n \quad (1)$$

where:

Σx = the sum of all measurements, and
 n = the number of measurements.

NOTE 4—Where all items of a category are included in a summation, the simple summation symbol Σ will be used. However, the strict mathematical statement of Eq 1 is:

$$\bar{x} = \sum_{i=1}^n (x_i)/n$$

where $x_1, x_2, \dots, x_n$ is the population of all n determinations which were made. Since no other constrictions are being made on the summation, the simpler statement of Eq 1 clearly shows the required operation.

5.2 Variance (s²):

5.2.1 Following directly from the definition of 3.2.9 (see Note 4 and Note 5):

$$s^2 = \Sigma(x_i - \bar{x})^2/(n - 1) \quad (2)$$

where:

x_i = an individual determination, and
 n = the number of determinations.

5.2.2 An alternative determination that can be readily handled with a calculator without first determining $\bar{x}$ is (see Note 4):

$$s^2 = [\Sigma(x^2) - (\Sigma x)^2/n]/(n - 1) \quad (3)$$

NOTE 5—To prevent significant errors in calculating s^2 do not round the sum of the squares of differences, $\Sigma(x_i - \bar{x})^2$ in Eq 2, nor the sum of the squares of the measurement, $\Sigma(x^2)$, and the square of the sum, $(\Sigma x)^2$, of Eq 3. Although these equations are algebraically identical, they may give slightly different results with large numbers or large summations.



computer because of greater round off errors from using Eq 3 instead of Eq 2.

NOTE 6—The numerators of Eq 2 and Eq 3 are often referred to as the “sum of squares,” meaning the sum of the squares of deviation. Using Eq 3, its effectiveness in being a measure of the degree of deviation of a series of values from each other can be seen by considering some simple lists. For example, 2, 3, 4 totals up to 9 for an average of 3. Summing the squares of the three values yields $4 + 9 + 16 = 29$. The square of the sum divided by the number of values is $9^2/3 = 81/3 = 27$. The sum of squares becomes $29 - 27 = 2$. If the list were 1, 3, 5, the sum would still be 9 and the term $(\Sigma x)^2/n$ would remain as 27. The summing of the squares of the three values, however, would yield $1 + 9 + 25 = 35$, and the sum of squares would now become $35 - 27 = 8$, giving a quantitative statement as to how much more deviation there is in the second set than the first. The sum of squares is insensitive to the level of the readings. Thus, if the list were 3, 4, 5, the summing would yield $9 + 16 + 25 = 50$ and $50 - 48 = 2$, resulting in the same difference as the first set as it should since the deviations are the same. If the list were 3, 3, 3, the $\Sigma(x^2)$ term would become 27 and the sum of squares would properly be zero since the $(\Sigma x)^2/n$ would also be 27 and the list shows no deviation.

5.2.3 *Estimate from Duplicate Determinations*—The difference between duplicate determinations can be used to estimate variance from $s^2 = \Delta^2/2$ where Δ is the difference. It is a special case of Eq 3, for then, if the two measurements are x_1 and x_2 :

$$\begin{aligned} s^2 &= [x_1^2 + x_2^2 - (x_1 + x_2)^2/2]/(2 - 1) \\ &= [2(x_1^2 + x_2^2) - (x_1^2 + 2x_1x_2 + x_2^2)]/2 \\ &= \frac{1}{2}(x_1^2 - 2x_1x_2 + x_2^2) = \frac{1}{2}(x_1 - x_2)^2 \end{aligned} \quad (4)$$

When there are K duplicates, the pooling of the individual estimates, as in 5.2.4, becomes (see Note 4):

$$s^2 = \Sigma \Delta^2/2K \quad (5)$$

If many duplicates are used, the degrees of freedom increases and the quality of the estimate of s^2 improves.

NOTE 7—The estimate from duplicates is particularly useful in production laboratories where routinely analyzed samples can be analyzed a second time to obtain a measure of precision under practical conditions.

5.2.4 *Pooled Estimate of Variance*—Where there is more than one set of similar repeat determinations, an improved overall estimate of variance may be made by pooling the individual estimates. Weight each individual variance, s_i^2 , by its degrees of freedom, sum the weighted variances, and divide by the total degrees of freedom:

$$\begin{aligned} s_p^2 &= [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 \\ &+ \dots + (n_k - 1)s_k^2]/[(n_1 - 1) \\ &+ (n_2 - 1) + \dots + (n_k - 1)] \end{aligned} \quad (6)$$

where:

- n_i = the number of measurements in each set, and therefore
- $(n_i - 1)$ = each degrees of freedom, and
- k = the number of sets.

5.2.4.1 In the special case of Eq 4, all of the $(n_i - 1)$ values are one, allowing for the summation of the $\Delta^2/2$ values with no weighting. If individual variances were not determined previously, the pooled estimate of variance can take the following form (see Note 4):

$$s_p^2 = \{\Sigma(x_i^2) - [(\Sigma x_1)^2/n_1 + (\Sigma x_2)^2/n_2 + \dots + (\Sigma x_k)^2/n_k]\}/(N - k) \quad (7)$$

$$+ \dots + (\Sigma x_k)^2/n_k]/(N - k)$$

where:

- x_i represents all readings, and
- N = the total number of readings.

NOTE 8—Pooling two or more estimates of variance is valid only if each set of analyses was obtained under similar or identical conditions with samples of similar composition and history. The pooling is valid only if variabilities are statistically the same.

5.3 *Standard Deviation (s)*—The estimate of standard deviation follows directly from variance as:

$$s = \sqrt{s^2} \quad (8)$$

5.3.1 A close estimate of s can be determined from the value of range, w , as defined in 3.2.13:

$$s_r = w/\sqrt{n'} = (x_h - x_l)/\sqrt{n'} \quad (9)$$

where:

- x_h = the highest measurement in a set,
- x_l = the lowest measurement in a set, and
- n' = number of measurements, limited from 4 through 12.

Reliable estimates from range by Eq 8 can be made only for sets of measurements from four through twelve. If extended beyond twelve measurements, the estimate will be low.

5.3.2 Repeat measurements, even when made on different days, might be biased because the second and subsequent values are expected to agree with the initial value. In cooperative analyses, a laboratory might make extra determinations and report only those that show good agreement. To overcome the possibility of such prejudiced results, an estimate of standard deviation may be calculated from single determinations made on pairs of samples having similar composition in a number of different laboratories or by a number of different analysts. For a single pair, determine the differences in measurement and pool as a special set (see Note 4):

$$s_d = \sqrt{[\Sigma(D^2) - (\Sigma D)^2/T]/2(T - 1)} \quad (10)$$

where:

- D = difference between values reported, and
- T = number of laboratories or analysts.

5.4 *Relative Standard Deviation (RSD)*:

$$RSD = 100s/\bar{x} \quad (11)$$

where:

- s = standard deviation estimate, and
- $\bar{x}$ = average.

5.4.1 A pooling of relative standard deviation can be done in the manner of Eq 5 after squaring the RSD values. The precaution of Note 8 applies with the exception that the concentration of the element being measured may vary, but only to the extent that there is no pattern of RSD varying with concentration. For example, low concentrations might show higher RSD values than high concentrations, making it inappropriate to pool RSD values over such a wide range of concentration. Using ν to represent RSD, the pooled estimate would be:





$$v = \sqrt{\frac{(n_1 - 1)v_1^2 + (n_2 - 1)v_2^2 + \dots + (n_k - 1)v_k^2}{(n_1 - 1) + (n_2 - 1) + \dots + (n_k - 1)}} \quad (12)$$

5.5 *Standard Error*—Determine from the differences between the assumed true concentrations of reference materials and the concentrations calculated from a calibration using (see Note 4):

$$SE = \sqrt{\Sigma(c_d - c_t)^2/f} \quad (13)$$

where:

c_d = a concentration determined from a calibration,

c_t = an accepted true concentration of the same material, and

f = the degrees of freedom in the observations.

If SE is calculated to determine how well the reference materials fit the curve they have defined, f is the number of observations minus the number of constants in the calibration curve. If SE is calculated with an independent set of reference materials, f equals the number of observations.

5.6 *Confidence Interval*—As discussed in 3.2.3, limits may be set by adding and subtracting from $\bar{x}$ the quantity $k\sigma/\sqrt{n}$ for $\bar{x} - k\sigma/\sqrt{n} \leq \mu \leq \bar{x} + k\sigma/\sqrt{n}$ where the value of k depends upon the confidence level desired. Specifically, for a 95 % confidence level, $k = 1.960$; for 99 %, $k = 2.576$; and for 90 %, $k = 1.645$. These are the “ t ” values which appear in Table 1 for infinite degrees of freedom which is used when the variability is not estimated but known.

5.6.1 In practice, the estimate, s , is known rather than σ . To reflect the uncertainty of substituting s for σ , a larger factor is used to determine the likely interval straddling μ . The practical

limit is then calculated by using $ts/\sqrt{n}$ for $k\sigma/\sqrt{n}$ permitting the statement of confidence interval as:

$$\bar{x} - ts/\sqrt{n} \leq \mu \leq \bar{x} + ts/\sqrt{n} \quad (14)$$

where the multiplier, t , is taken from a Student’s t -table, an extract of which is shown in Table 1. The square root of n takes into account that the limits apply to the average of n determinations. The “ t ” values have been computed on the assumption that the errors follow the normal error distribution curve and that the measurements are independent.

5.6.2 Eq 13 can be restated as $\bar{x} - \mu \leq ts/\sqrt{n}$ and $\mu - \bar{x} \leq ts/\sqrt{n}$. When the “ t ” values become critical these expressions become equalities, $\bar{x} - \mu = ts/\sqrt{n}$ and $\mu - \bar{x} = ts/\sqrt{n}$, making the critical “ t ” value become:

$$t = |\bar{x} - \mu| \sqrt{n}/s \quad (15)$$

NOTE 9—The t and F tables show the probability, decimally, for measurements not to exceed certain values (see Tables 1-3). These translate directly to the confidence level by subtracting the probability from 1 and multiplying by 100. Thus, if the probability for a measurement not to exceed a certain value is 0.05, the confidence level that the measurement will fall within the range indicated by the critical value is $(1 - 0.05)100 = 95\%$.

5.7 *Tests for Outliers*—A detailed test for outliers is described in Practice E 178. Some commonly used procedures are given below:

NOTE 10—An apparent outlying value may be merely an extreme example of random variability, in which case it should be retained with other data, or it may be the result of gross deviation from the prescribed analytical procedure, or an error in calculating or recording the analysis. The reason for aberrant values may warrant investigation.

5.7.1 For four or more determinations, estimate the standard deviation and then repeat the calculation with the suspect value eliminated. If the second estimate is smaller by a factor of two or more, the suspect value may be excluded. For this test, it may be particularly useful to use the estimate of s , by range of Eq 8.

5.7.2 For triplicate determinations, list the values in increasing or decreasing order as R_1 , R_2 , and R_3 , with R_1 being the suspect value. Calculate as follows:

$$\text{Test ratio} = (R_1 - R_2)/(R_1 - R_3), \text{ provided } R_2 \neq R_3 \quad (16)$$

If the test ratio is larger than 0.941, the suspect value may be rejected with a 95 % confidence that it is an outlier. When the readings are not precise enough to show a difference between R_2 and R_3 , the test may still be done by making one reading higher by 5 in the next decimal place and the other lower by that amount. For example, if the two readings are 1.23, one can be made 1.235, and the other 1.225.

NOTE 11—The test in 5.7.2 is a special case of the Dixon Criteria. Practice E 178 shows how the Dixon test can be applied to up to 30 observations and lists significant test values for 90 %, 95 %, and 99 % confidence levels.

5.7.3 A more reliable test can be made if there is a good record of the standard deviation or relative standard deviation expected for the level of readings or determinations being made. If pooling was used to make the estimate of s or RSD it should represent an estimation based on a relatively high degrees of freedom, permitting a favorable use of Student’s

TABLE 1 Student’s t -Table^A

Degrees of Freedom	Probability Level ^B		
	0.10	0.05	0.01
1	6.314	12.706	63.657
2	2.920	4.303	9.925
3	2.353	3.182	5.841
4	2.132	2.776	4.604
5	2.015	2.571	4.032
6	1.943	2.447	3.707
7	1.895	2.365	3.499
8	1.860	2.306	3.355
9	1.833	2.262	3.250
10	1.812	2.228	3.169
12	1.782	2.179	3.055
14	1.761	2.145	2.977
16	1.746	2.120	2.921
18	1.734	2.101	2.878
20	1.725	2.086	2.845
25	1.708	2.060	2.787
30	1.696	2.042	2.750
∞	1.6448	1.9600	2.5758

^A Table available in any standard publication of statistical tables, usually entitled the t -Test for Significance or Distribution of t (Two-Tailed Test). Credit usually given to Fisher, R. A., *Statistical Methods for Research Workers*, published by Oliver and Boyd, Edinburgh, Scotland, 1925–1950.

^B Probability, p , stated decimally, that values will exceed a mean by the stated “ t ” factor times the standard deviation which was estimated at the listed degrees of freedom. For the probability, or confidence level, that a measurement will not exceed this quantity, use $(1 - p) 100$. Thus, for a value for p of 0.05, the confidence level is $(1 - 0.05) 100 = 95\%$. The table can be converted to a One-tail Test by dividing the probability level by 2, such as is required for determining detection as described in 6.2.2.

TABLE 2 *F* Test for Equality of Variances, 0.05 Probability^A

Degrees of Freedom Denominator	Degrees of Freedom of Numerator, the Greater Mean Square												
	1	2	3	4	5	6	7	8	10	12	16	20	30
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.78	8.74	8.69	8.66	8.62
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.84	5.80	5.74
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.60	4.56	4.50
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.92	3.87	3.81
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.63	3.57	3.49	3.44	3.38
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.34	3.28	3.20	3.15	3.08
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.97	2.91	2.82	2.77	2.70
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.76	2.69	2.60	2.54	2.46
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.60	2.53	2.44	2.39	2.31
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.33	2.28	2.20
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.35	2.28	2.18	2.12	2.04
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.24	2.16	2.06	2.00	1.92
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.16	2.09	1.99	1.93	1.84

^A Extract of tables for the *F* test which are available in any standard publication of statistical tables. This specific extract is from Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, The Iowa State University Press, Ames, IA, Sixth Edition, 1967. The first column shows the degrees of freedom of the denominator, which would be the lesser of the mean squares being compared.

TABLE 3 *F* Test for Equality of Variances, 0.01 Probability^A

Degrees of Freedom Denominator	Degrees of Freedom of Numerator, the Greater Mean Square												
	1	2	3	4	5	6	7	8	10	12	16	20	30
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.23	27.05	26.83	26.69	26.50
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.54	14.37	14.15	14.02	13.83
5	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.29	10.05	9.89	9.68	9.55	9.38
6	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.87	7.72	7.52	7.39	7.23
7	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.62	6.47	6.27	6.15	5.98
8	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.82	5.67	5.48	5.36	5.20
10	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.85	4.71	4.52	4.41	4.25
12	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.30	4.16	3.98	3.86	3.70
14	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	3.94	3.80	3.62	3.51	3.34
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.69	3.55	3.37	3.25	3.10
20	8.10	5.85	4.94	4.43	4.10	3.87	3.71	3.56	3.37	3.23	3.05	2.94	2.77
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.13	2.99	2.81	2.70	2.54
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	2.98	2.84	2.66	2.55	2.38

^A Extract of tables for the *F* test which are available in any standard publication of statistical tables. This specific extract is from Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, The Iowa State University Press, Ames, IA, Sixth Edition, 1967. The first column shows the degrees of freedom of the denominator, which would be the lesser of the mean squares being compared.

t-Table (see Table 1). Calculate the average reading and use Eq 13 to determine a confidence interval that encompasses that average. If a suspect value does not fit within these limits, exclude it.

5.7.4 For duplicate determinations, if the values seem to be far apart and an estimate of standard deviation has already been established, a third value should be obtained if the difference between the two readings is greater than 2.8s. Subsequent testing, as in 5.7.2 or 5.7.3, would be appropriate to determine if one reading should be rejected.

6. Applications

6.1 *Testing for Bias*—Bias can be detected in various ways.

6.1.1 *Analysis of Variance*—A series of single measurements on a suite of samples can be subjected to an analysis of variance to determine the precision of the method and to detect bias in different sets of the running of the suite of samples. The sets may be the results determined by different methods, different laboratories, or different analysts. It is required that

any one grouping of samples be close in composition. Details on the analysis of variance appear in Snedecor and Cochran.⁴

6.1.1.1 Obtain the differences between actual measurements and the assumed correct values in order to put the values into a common population. Call the differences remainders or “*r*” values. Prepare a table of “*r*” values with samples in lines and sets in columns. Table 4 contains a general example of a tabulation of “*r*” values for *n* samples and *m* sets with summations under each column for Σr , $(\Sigma r)^2$ and $\Sigma(r^2)$ and with the grand totals of these.

6.1.1.2 Set up a table of comparisons of variance as shown in Table 5.

NOTE 12—The sum of squares for within samples could also be obtained by pooling the variances of the individual sets. The $B/n - G^2/mn$ used for the sum of squares between sets is equivalent to using the averages of each set, $\Sigma(\bar{x})^2 - (\Sigma \bar{x})^2/m$, when that sum is multiplied by the number of samples, *n*, to reflect that this measure of deviation is for an average of *n* readings instead of for single readings.

⁴ Snedecor, G. W., and Cochran, W. G., *Statistical Methods*, The Iowa State University Press, Ames, IA, Sixth Edition, 1967, pp. 258–269.