ASHRAE Guideline 2-1986 (RA 96)
(Reaffirmation of ASHRAE Guideline 2-1986
with minor editorial changes)

ASHRAE® GUIDELINE

Engineering Analysis of Experimental Data

Approved by the ASHRAE Standards Committee September 23, 1986; approved by the ASHRAE Board of Directors for publication December 12, 1986; and reaffirmed on June 14, 1990, and February 22, 1996.

ASHRAE Guidelines are updated on a five-year cycle; the date following the Guideline number is the year of ASHRAE Board of Directors approval. The designation "RA" followed by a date is the year of reaffirmation. The latest copies may be purchased from ASHRAE Customer Service, 1791 Tullie Circle, NE, Atlanta, GA 30329, (404) 636-8400.

ISSN 1049-894X

AMERICAN SOCIETY OF HEATING, REFRIGERATING AND AIR-CONDITIONING ENGINEERS, INC.
1791 Tullie Circle, NE • Atlanta, GA 30329

ASHRAE STANDARDS PROJECT COMMITTEE GPC 2-1986

ASHRAE STANDARDS COMMITTEE 1995-96

SPECIAL NOTE

This Guideline was developed under the auspices of the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). ASHRAE Guidelines are developed under a review process, identifying a guideline for the design, testing, application, or evaluation of a specific product, concept, or practice. As a guideline it is not definitive but encompasses areas where there may be a variety of approaches, none of which must be precisely correct. ASHRAE Guidelines are written to assist professionals in the area of concern and expertise of ASHRAE's Technical Committees and Task Groups.

ASHRAE Guidelines are prepared by project committees appointed specifically for he purpose of writing Guidelines. The project committee chair and vice-chair must be members of the ASHRAE; while other members of the project committee may or may not be ASHRAE members, all must be technically qualified in the subject area of the Guideline.

Development of ASHRAE Guidelines follows procedures similar to those for ASHRAE Standards except that (a) committee balance is desired but not required, (b) and effort is made to achieve consensus but consensus si not required, (c) guidelines are not appealable, and (d) guidelines are not submitted to ANSI for approval.

The Manager of Standards of ASHRAE should be contacted for:

a. interpretation of the contents of this Standard,
b. participation in the next review of the Standard,
c. offering constructive criticism for improving the Standard,
d. permission to reprint portions of the Standard.

ASHRAE INDUSTRIAL ADVERTISING POLICY ON STANDARDS

ASHRAE Standards and Guidelines are established to assist Industry and the public by offering a uniform method of testing for rating purposes, by suggesting safe practices in designing and installing equipment, by providing proper definitions of this equipment, and by providing other information that may serve to guide the industry. The creation of ASHRAE Standards and Guidelines is determined by the need for them, and conformance to them is completely voluntary.

In referring to this Standard or Guideline and in marking of equipment and in advertising, no claim shall be made, either stated or implied, that the product has been approved by ASHRAE.

DISCLAIMER

ASHRAE uses its best efforts to promulgate Standards and Guidelines for the benefit of the public in light of available information and accepted industry practices. However, ASHRAE does not guarantee, certify, or assure the safety or performance of any products, components, or systems tested, installed, or operated in accordance with ASHRAE's Standards or Guidelines or that any tests conducted under its Standards or Guidelines will be nonhazardous or free from risk.

CONTENTS

ASHRAE Guideline 2-1986 (RA96) Engineering Analysis of Experimental Data

This foreword is not part of this guideline but is included for information purposes only.

FOREWORD

The subject matter covered by this guideline was originally contained in ASHRAE Standard 41.5-75, Standard Measurement Guide: Engineering Analysis of Experimental Data. SPC 41.5-75R was formed to revise this standard, and a draft of ANSI/ASHRAE Standard 41.5-75R was submitted for public review in May 1984. Comments obtained during the public review suggested that the material contained in the standard was more appropriate as a guideline rather than a standard. SPC 41.5-75R agreed and unanimously voted to publish the document as a guideline. This guideline was then approved by the Standards Committee on October 25, 1986; the Board of Directors on December 12, 1986; and by the American National Standards Institute on February 18, 1987. The guideline was subsequently reaffirmed in 1990 and again in 1996.

1. PURPOSE

The purpose of this guideline is to provide guidelines for reporting uncertainty in results of experimental data obtained from the testing of heating, refrigerating, and air-conditioning equipment.

2. SCOPE

Appropriate terms are defined and recommended procedures are given for applying basic statistical methods to experimental data.

3. DEFINITIONS AND TERMINOLOGY

confidence level: probability that a stated interval will include the true mean.

confidence limits (used for multisample data) and ***uncertainty interval*** (used for single-sample data): that range of values that can be expected to include the true value with a stated probability. For example, a statement that the 90% confidence limit is 5 to 8 means that there is 90% probability that the interval between 5 and 8 will contain the true value.

correlation coefficient (coefficient of determination): a measure of the linear relationship between two quantitative variables.

data: the information obtained by experimental means, assumed to be in numerical form; recorded values of the variables; readings.

deviation: the difference between a single result and the mean of many results.

- ***mean deviation***: the sum of the absolute values of the deviations divided by their number.
- ***residual standard deviation***: the standard deviation compiled from the residuals from the fitted curve.
- ***standard deviation***: the square root of the mean of the squares of the deviations. (Adjustment is necessary for a small number of measurements, see 6.4.2.)

error: the difference between the true value of the quantity measured and the observed value. All experimental errors can be classified as one of two types: systematic (accuracy) error or random (precision) error.[1] See Example 1 in the appendix.

- ***fixed error***: same as systematic error.
- ***precision error***: same as random error.
- ***random error***: an error that causes readings to take random values on either side of some mean value. Measurements may be precise or imprecise depending on how well an instrument can reproduce subsequent readings of an unchanged input.
- ***systematic error***: an error that persists and cannot be considered as due entirely to chance. Systematic error can be corrected through calibration.
- ***uncertainty***: an estimated value for the error, i.e., what we think an error would be if we could and did measure it by calibration. Although uncertainty may be the result of both accuracy and precision errors, only precision error can be treated by statistical methods, and in this report, uncertainty will always be analyzed as a precision error.

interval range: the range between the largest and smallest values.

mean: the sum of measurement values divided by the number of measurements. It is considered the "best" approximation of the true value.

population: a collection of items, representing the same quantity and connoting completeness; the entire group of items being studied whether they be instrument readings, test results, points on a curve, or parts from a factory production lot.

propagation of uncertainty: the degree to which the uncertainties in the values of the parameters affect the uncertainty in the result.

result: the value obtained by making corrections to or calculations with the data.

sample: a portion or limited number of items of a population; a set of values, experimentally obtained, that represents a sample of all possible readings that could be taken.

- ***multisample***: repeated measurement of a fixed quantity using altered test conditions, such as different observers or different instrumentation or both; experiments in which uncertainties are evaluated by sufficient repetition using enough observers and enough diverse instruments so that the reliability of the results can be ensured by the use of statistics. Merely taking repeated readings with the same procedure and equipment does not provide multisample results.
- ***single sample***: a single reading or succession of readings taken at the same or different times but under identical conditions. Many experiments that appear to be multisample are actually, in part, single-sample experiments. If the same instrument is used for a set of observations, some error that is inherent in the instrument will be sampled only once, no matter how many times each reading is repeated. In single-sample experiments, the statement of uncertainty or reliability will be based in part on estimates of possible systematic errors based on judgment or experience, since statistics cannot be applied to

[1] "Accuracy" and "precision" are often used to distinguish between systematic and random errors. A measurement with small systematic errors is said to have high accuracy. A measurement with small random errors is said to have high precision.

all of the errors.

variable: the basic quantity observed.

variance: the square of the standard deviation.

4. GENERAL CONSIDERATIONS IN UNCERTAINTY ANALYSIS AND EXPERIMENTAL PLANNING

4.1 All measurements or results should be given in three parts:

- Best value
- ± Confidence limits or uncertainty interval
- Specified probability

Reporting of experimental data should first provide an estimate of the value of the dependent variable at a point and, coupled with this, an interval, based on a selected degree of confidence, within which the true value is expected to lie.

4.2 Generalized Experimental Procedure

4.2.1 Establish the quantities that must be measured (temperature, pressure, flow, etc.).

4.2.2 Set up the formulas for calculation of the desired result.

4.2.3 Decide the maximum uncertainty tolerable in the result.

4.2.4 Select instrumentation. The calculation of uncertainty in the result requires knowledge of individual uncertainties. Therefore, proposed instrumentation would have to be known at this point.

4.2.5 Estimate the uncertainty in a single measurement of each parameter.

4.2.6 Calculate the uncertainty in the anticipated result before the experiments are conducted by properly combining the estimated uncertainties in the value of each parameter as described in Section 5. An uncertainty analysis is performed on each measurement technique, taking into account the estimated uncertainties of the instruments that will actually be used.

4.2.7 Identify as critical those measurements whose uncertainties must be reduced. Also determine the amounts of reduction required if the estimated uncertainty in the results is greater than the maximum allowed. Compare the different measurement techniques for the critical measurement on the basis of cost, availability of instrumentation, ease of data collection, and calculated uncertainty. Select the measurement technique that provides the best compromise between small uncertainty and high cost.

4.2.8 Collect a few data points and make a preliminary analysis of these data.

4.2.9 Modify the experimental apparatus or procedure, or both, based on the results of 4.2.8.

4.2.10 Collect the experimental data and analyze the results following procedures in Sections 4.3, 5, and 6.

4.2.11 Report results as specified in 4.1.

4.3 General Considerations in Data Analysis

4.3.1 Examine the data for consistency. If any data points appear inconsistent with physical reality (e.g., heat is added to a container of water and the measured temperature drops), the related experimental procedure should be investigated for gross mistakes or miscalculations.

4.3.2 Perform a statistical analysis of data where appropriate (Section 6). Discard erroneous data if there is sound basis for rejection (see 6.5).

4.3.3 Estimate the uncertainty in the results (Section 5).

4.4 Uncertainties in multisample data are treated in Section 6.

4.5 Uncertainties in single-sample data are treated in Section 7.

5. PROPAGATION OF UNCERTAINTIES

5.1 In experiments, the measurements eventually must be used to calculate either the desired result or a variable to be used in such calculation.

5.1.1 Uncertainties can be described either in absolute terms (i.e., 5 volts ±0.5 volts, $P = 83\%$) or in percentage terms (i.e., 5 volts ±10%, $P = 83\%$).

5.1.2 Generally, as described more precisely below, uncertain quantities that are to be added have their uncertainties expressed in absolute terms; uncertain quantities that are to be multiplied, divided, or exponentiated have their uncertainties expressed as percentages.

5.2 Approximations

5.2.1 Select the measurement or variable with the greatest percent uncertainty and assign this uncertainty to the result. This can be a simple and satisfactory procedure when the most uncertain measurement or variable has an uncertainty that is four or more times the next largest uncertainty. See Example 2 in the appendix.

5.2.2 The uncertainty in the result could be obtained by combining all the uncertainties in the most detrimental way (i.e., by adding them). This method should be used where there are only a few measurements or variables and the uncertainties are small. See Example 2 in the appendix.

5.3 With the above exception, the uncertainty in a result is not closely approximated by the sum of the absolute or percentage uncertainties of the values of the parameters. This would imply high odds that the direction of all uncertainties might simultaneously be the same. Since the concept of uncertainty involves random deviation about a mean, some deviations would be positive while others would be negative. Therefore, a series of measurements with relatively large uncertainties could produce a result with an uncertainty not much larger than the uncertainty of the most uncertain measurement. Subsection 5.4 sets forth the technique whereby this result is achieved.

5.4 Recommended Procedure

5.4.1 The uncertainty in each measurement or variable is described by specifying the expected value for the variable followed by the absolute uncertainty followed by the confidence level.

$$v = m \pm w;\ P \text{ percent}$$

where

v = the variable,

m = its best value,

w = uncertainty, and

P = confidence level.

For example: pressure = 50.2 psia ± 0.5; 95% states that the best value for the pressure is believed to be 50.2 psia with a 95% probability that the true value lies within ± 0.5 psia of this value.

5.4.2 Where instrument accuracy is stated in percent of full scale, the maximum expected absolute uncertainty of any reading is the product of the stated percent and the value of the highest scale reading. This absolute uncertainty must be applied either directly or as a percentage of the expected scale reading.

Instrumental accuracies that are so stated are generally at a confidence level of 95%. These 95% values may be multiplied by the following factors for other confidence levels:

Confidence Level, %	Factor	For Information: Areas of the normal curve at x/σ
50	0.34	0.67
60	0.43	0.85
70	0.53	1.04
80	0.65	1.28
90	0.84	1.64
95	1.0	1.96

5.4.3 The rule for combining uncertainties of the same confidence level for the mathematical operations of addition and subtraction of quantities is that the squares of the absolute value of the uncertainty for each independently measured quantity are added. The square root of their sum is taken (root mean square, or rms) and is expressed as a percent of the sum or difference of the quantities.

The rule for combining uncertainties for the arithmetical operations of multiplication, division, and exponentiation is that the uncertainty in the test result is the root mean square (rms) of the fractional uncertainty of each of the independently measured test variables.

For purposes of illustration, assume measurements A and B of two independent variables are made and that an uncertainty of $\pm a$ exists in the measurement of A and an uncertainty of $\pm b$ exists in the measurement of B. If w_r represents the resultant uncertainty after various arithmetic operations on numbers that contain the uncertainties (in the units of the measurements), then:

For addition, $(A \pm a) + (B \pm b)$

$$w_r = (a^2 + b^2)^{0.5}$$

For subtraction, $(A \pm a) - (B \pm b)$

$$w_r = (a^2 + b^2)^{0.5}$$

For multiplication, $(A \pm a)\,(B \pm b)$

$$w_r = [(a/A)^2 + (b/B)^2]^{0.5}$$

For division, $(B \pm b) \div (A \pm a)$

$$w_r = B/A\ [(a/A)^2 + (b/B)^2]^{0.5}$$

5.4.4 In general, if the uncertainties in the independently measured variables are all given with the same confidence level, then the uncertainty in the result having this confidence level is

$$w_r = [(\partial R/\partial v_1\, w_1)^2 + (\partial R/\partial v_2\, w_2)^2 + \ldots + (\partial R/\partial v_n\, w_n)^2]^{0.5}$$

where

R is a given function of the independent variables $v_1, v_2, \ldots, v_n$ or

$R = R(v_1, v_2, \ldots v_n)$

with w_r the uncertainty in the result and $w_1, w_2, \ldots w_n$ the uncertainties in the independently measured variables. See Examples 2 and 3 in the appendix.

5.4.5 Particular notice should be given to the fact that the uncertainty propagation in the result depends on the squares of the uncertainties in the independent variables. This means that if the uncertainty in one variable is significantly larger than the uncertainties in the other variables, say by a factor of 5 or 10, then it is the largest uncertainty that predominates, and the other may probably be neglected. To illustrate, suppose there are three variables with an uncertainty of magnitude 1 and one variable with an uncertainty of magnitude 5. The uncertainty in the result would be

$$(5^2 + 1^2 + 1^2 + 1^2)^{0.5} = \sqrt{28} = 5.29$$

The importance of these brief remarks concerning the relative magnitude of uncertainties is evident when one considers the design of an experiment, the procurement of instrumentation, etc. Very little is gained by trying to reduce the "small" uncertainties. Because of the square propagation, it is the "large" ones that predominate. Any improvement in the overall experimental result must be achieved by improving the instrumentation or experimental technique connected with these relatively large uncertainties.

5.4.6 Summary of Recommended Procedure

5.4.6.1 Describe the uncertainty (w) in each variable (v).

v = best (mean) value $\pm w$

(with P percent [95% recommended] confidence)

5.4.6.2 Compute the uncertainty in each result as

$$w_r = \sqrt{(\partial R/\partial v_1 w_1)^2 + (\partial R/\partial v_2 w_2)^2 + \ldots + (\partial R/\partial v_n w_n)^2}$$

5.4.6.3 Report at least $\pm w_r$ at the chosen confidence level.

5.4.6.4 The value of w_r will be based on essentially the same confidence level as the uncertainties in the variables.

6. TREATMENT OF MULTISAMPLE DATA

6.1 General Considerations

Statistical techniques provide methods for dealing economically with a representative sample of a large number of items. For example, the degree of precision of a mean of observed values or of plotted results can always be improved by increasing the number of observations or test points. However, the same improvement in precision can be made at much less cost, in many cases, by statistical design of the experiment with fewer data points. Again, a statistical analysis of test results based on a proper design may give the same information with a few tests that could otherwise be obtained only by

a large number of tests of conventional type, if there are two or more variables that influence the process.

For the majority of engineering cases, it is impractical and too costly to perform a true multisample experiment. However, repeated measurements with the same instrumentation and the same observer may provide a fair approximation of a multisample experiment. Therefore, the following recommended procedures are based strictly on sample size:

More than 20:	Multisample, Subsection 6.3 procedure
2 to 20:	Small Sample, Subsection 6.4 procedure
1:	Single Sample, Section 7.0 procedure

When a number of observations are experimentally obtained, the data values are dispersed or scattered about the mean. The worthiness of data of this sort may be evaluated through application of certain statistical rules. The distribution of the values in the sample, or in the entire population, may be compared with some expected probability distribution. One interpretation of probability is linked to the frequency with which a certain phenomenon occurs after a large number of tries. There are several recognized curves for probability distribution, but of the most use here is the normal error distribution, also called the normal frequency or Gaussian probability curve.

While the Gaussian distribution does not fit all experimental data sets, it is generally assumed for the following reasons:

- least squares estimators are "best" for Gaussian data, and they are easy to compute;
- the theory behind classical statistical tests (*F*-tests, *t*-tests, etc.) requires the assumption of a Gaussian distribution;
- such classical tests have precise levels only for Gaussian data; and
- most of these tests are fairly good even with respect to "stretched-tailed" data.

6.2 The Gaussian or Normal Distribution

6.2.1 When data abide by Gaussian or normal distribution rules, plus or minus errors are equally probable, and small errors are more probable than large errors. However, there is no real limit to the magnitude of large errors. The equation for such a distribution, assuming an infinite population, may be written as follows:

$$y_x = h/\sqrt{\pi}\, e^{-h^2 (x-m)^2}$$

where

x = dependent variable or output reading;
y_x = the frequency of occurrence of the reading, x, or the probability of its occurrence;
m = the mean value of the population; and
h = a constant inversely proportional to the standard deviation of the population ($h = 1/\sqrt{2}\sigma$).

Figure 1 is a plot of this function for two different values of h. It is seen that the distribution represented by the dashed curve indicates a greater number of small errors and a correspondingly smaller number of large errors. Data represented by the dashed curve are said to be more precise than those represented by the solid curve. The curves are bell-shaped, and

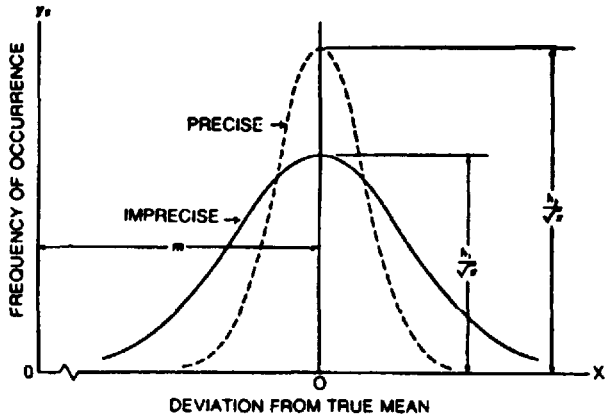

Figure 1 Normal distribution curve.

three common measures of central tendency (the mean, the mode, and the median) all coincide. The mode is the value that occurs most frequently, and the median is the middle value.

6.2.2 One may ask how it is known that the assumptions pertaining to the derivation of the normal error distribution apply to experimental data. Actually, normality is a condition that, rigorously, may never occur. For example, consider repeated readings with a pressure gauge at 4 psig. If the standard deviation of the instrument is 2 psi, then there is a small but finite probability that this gauge will indicate 10 psi ($+3\sigma$). However, there is no possibility for the gauge to indicate -2 psi (-3σ) because of the pin that stops the indicator dial at 0 psig. Therefore, this instrument, or any instrument with a zero or maximum point, cannot deviate in a perfectly normal manner. Nevertheless, the normal distribution is a good approximation to the deviation patterns of many instruments.

Although experimental data may approximate the normal distribution, they seldom, if ever, fit the bell-shaped curve exactly. Goodness of fit, therefore, becomes a legitimate question, and various methods have been devised to evaluate this factor. The simplest method is to plot the data on probability paper on which a normal distribution plots as a straight line that passes through the point (0, 50) (Figure 2). Also shown in Figure 2 are two types of skewed curves and two types of flat curves. Skewed curves have a different slope on either side of the maximum point, while flatness refers to curves that peak either too sharply or too bluntly. Now, the question arises as to how straight and how close to the (0, 50) point one should pass if the distribution is to be considered normal. Usually, considerable flatness can be accepted without much concern. However, an imprecise instrument that has a badly skewed distribution has very likely malfunctioned and should either be repaired or discarded. Thus, as a matter of experimental verification, the Gaussian distribution is believed to represent the distribution of random errors in an adequate manner for a properly controlled experiment. See Example 4 in the appendix.

6.2.3 Best Value at a Given Input. When an output x is measured many times at a given input, the mean value of x is simply

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

where

x_i = the value of the ith observation,
n = the number of observations in the sample.

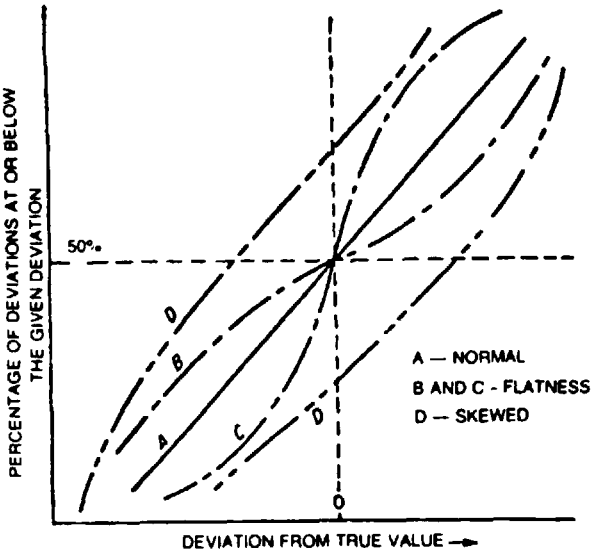

Figure 2 Normal and other distributions on probability coordinates.

This arithmetic average is the best representation of the given set of x_i's. Note that when the estimated best value of $\bar{x}$ is taken as 2, the sum of the squares of the deviation of the data from their estimate is a minimum. (This is essentially the least squares principle.) However, while $\bar{x}$ represents an unbiased estimate of the true arithmetic mean of all possible values of x, there is no assurance that $\bar{x}$ is the true value. Good agreement (i.e., high precision) in replication does not imply that $\bar{x}$ is close to the true value (i.e., high accuracy). Nevertheless, from any viewpoint, the best estimate of the value of an unknown parameter at a given input is the average of the available measurements.

6.2.4 Deviation. The deviation, d, for each reading is defined by:

$$d_i = x_i - \bar{x}$$

The average of the deviations of all the readings is zero since

$$\bar{d}_i = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x}) = \bar{x} - \frac{1}{n}(n\bar{x}) = 0$$

6.2.5 Mean Deviation. The average of the absolute values of the deviations, the mean deviation, is given by

$$\bar{d}_i = \frac{1}{n}\sum_{i=1}^{n} |d_i| = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$$

Note that the numerical values of the deviations are added without regard to algebraic sign, and that this quantity is not zero unless all the individual deviations are zero.

As stated previously, the actual value by which a result is in error is never known. There are, however, various ways of estimating error. Mean deviation (also known as probable error) is one. This quantity is not, however, the most popularly used error estimate.

6.2.6 Standard Deviation. The standard deviation (σ) is defined for very large n by

$$\sigma = \left[\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2\right]^{0.5} = \frac{\sqrt{(d_1)^2 + (d_2)^2 + \ldots + (d_n)^2}}{n}$$

For small n, say less than 20, the division is by $n-1$ rather than n, and the symbol σ' is used to indicate an estimate of sigma. The standard deviation is probably the most widely used indicator of the precision error of an instrument system. For a normal distribution, the standard deviation has the following meanings:

68.3% of the data will be within ± 1σ of the mean

95.5% of the data will be within ± 2σ of the mean

99.7% of the data will be within ± 3σ of the mean

The standard deviation indicates the degree of dispersion of the values about the mean. Figure 3 is a graphical statement of the deviation of all values from the mean or arithmetical average. The area under the curve, if measured symmetrically on both sides of the mean or zero-deviation value, can be interpreted to represent the percentage of items (or observations or quantities) falling within the limits shown by the abscissas or the probability that the original values will fall within these limits.

When the deviations from the mean value are examined and the average deviation and standard (rms) deviation σ are known, the usefulness of this distribution curve becomes apparent. This is because of certain conventional practices and available tables. For example, 68.3% of all readings should lie within ±σ of the average, and 99.7% should lie within ±3σ, where σ is the standard deviation or rms error.

Other arbitrary percentile selections could have been made and the results obtained with the aid of Table 1. Some statisticians give the 50%, 90%, and 99% values. One-half of the values lie outside of ±0.675σ, and the deviation at this point is called the probable rms error.

6.2.7 A very important question has not yet been answered: How good (or precise) is the arithmetic mean value that is taken as the best estimate of the true value of a set of readings?

The question may be reworded, "Is it not possible to obtain more than one sample of a given population?" If this is so, should one not expect that each sample would yield a slightly different average or mean?

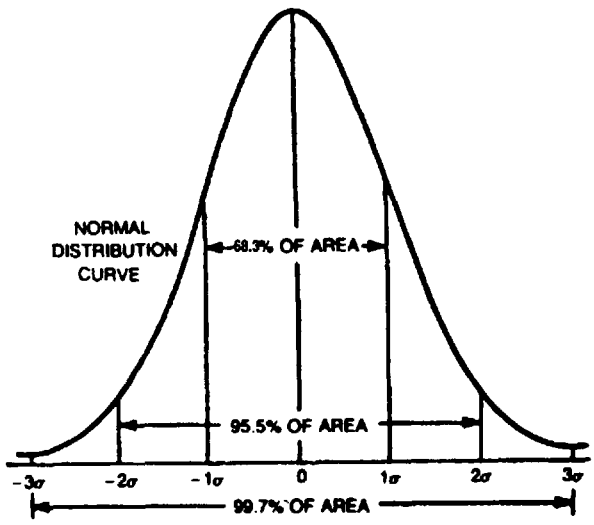

Figure 3 Normal probability curve of deviation or error in terms of standard deviation σ.

To obtain an experimental answer to this question, it would be necessary to repeat the set of measurements and find a new arithmetic mean. In general, this new arithmetic mean would differ from the previous value, and the problem would remain unresolved until a large number of sets of data were collected. It would then be known how well the mean of a single set approximated the mean that would be obtained with a large number of sets. The mean value of a large number of sets is presumably the true value. Consequently, the standard deviation of the mean of a single set of data from this true value is desired. The result is

$$\bar{\sigma} = \sigma/\sqrt{n}$$

where

$\bar{\sigma}$ = standard-deviation-of-the-mean value,
σ = standard deviation of the set of measurements, and
n = number of measurements in the set.

Thus, we can state with 99.7% confidence that the true value will not differ by more than $\pm 3\bar{\sigma}$ from the mean of the sample, etc.

6.3 Recommended Procedure for Multisample (More Than 20) Data with Normal Distribution Assumed. See Examples 5 and 6 in the appendix.

6.3.1 Best Value. The best value of a parameter is the mean value or a least squares adjusted value.

6.3.2 Confidence Limits. Having decided on the best available value of x, its worth as an estimate of the true value of x must be determined by constructing the limits of the confidence interval. The recommended method is to take the unbiased estimate of the parameter as given by the mean of the readings and say that the population lies within $\pm Z$ standard deviations from the observed mean.

The formula for the confidence limits is:

$$\text{Confidence Limits} = \bar{x} \pm (Z\sigma/\sqrt{n})$$

where

n = the sample size or the number of readings taken,
$\bar{x}$ = the average of the readings, and
σ = the standard deviation of the population.

For example, if we let $Z = 2$, there is a 95.5% probability that the correct value lies in the interval between $\bar{x} - (2\sigma/\sqrt{n})$ and $\bar{x} + (2\sigma/\sqrt{n})$.

It should be noted that Z is equivalent to Student's "t" (defined in 6.4.3) for infinite degrees of freedom (or number of data points). It will be more correct to use Student's "t" value for Z when there are a reasonably small number of data points.

6.4 Method of Estimating Confidence in Small Sample Size (2 to 20). See Examples 7 and 10 in the appendix.

6.4.1 Mean Value. In many circumstances the engineer will not be able to collect as many data points as might be desired, and only an approximation to the Gaussian distribution will be obtained. Generally, it is necessary to have a minimum of 20 measurements (100 is preferable) in order to obtain reliable estimates of standard deviation and general validity of the data For small sets of data, a mean value is computed as before:

$$\bar{x}' = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Note that a prime designation has been added, with $\bar{x}'$ referred to as the estimated mean of the population. Because of the small number of readings, the estimated mean usually will not be exactly equal to the true mean or correct value of the measured variable.

6.4.2 Standard Deviation. The following relation is then used as the best estimate for the standard deviation:

$$\sigma' = \left[\frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x}')^2\right]^{0.5}$$

Note that the factor $n-1$ is used in the denominator instead of n in an attempt to compensate for using the estimated mean instead of the true mean; however, some bias always remains, and σ', the obtainable, does not equal σ, the desired.

6.4.3 "Student" *t*-Distribution. Recognizing this deficiency, a method was developed by W.S. Gosset (writing under the pseudonym "Student") in 1907 by which confidence limits could be based on the standard deviation σ of the small sample. He introduced the "Student" t-distribution, which has been tabulated as a constant depending on degrees of freedom and on the desired degree of confidence. (See Table 2.) Degrees of freedom can be defined in general as the number of observations minus the number of different quantities estimated. In 10 readings of a single variable, such as temperature or pressure, the number of degrees of freedom equals 9. In 10 values for the area of a rectangle, the number of degrees of freedom equals 8, since there are two quantities, length and width, whose values determine the value of A.

On the basis of the t-distribution, a confidence interval for the mean value of x is:

TABLE 1 Summary of Error Estimates Based on Normal Distribution

Name of Error	Symbol	Value in Terms of σ	Percent Certainty	Probability That a Single Value Will Be Greater
Probable error (also mean deviation)	E_p	0.6745σ	50	1 in 2
Standard Deviation	σ	σ	68.3	1 in 3 (approx.)
90% error	E_{90}	1.6449σ	90	1 in 10
Two sigma error	2σ	2σ	95	1 in 20
Three sigma error	3σ	3σ	99.7	1 in 370

TABLE 2 Student's *t*-Values

F, Degrees of Freedom	*P*, Probability of Obtaining a Given Value of *t* or a Larger One (Two-Sided)							
	0.50	**0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.001**	**0.0005**
1	1.000	3.078	6.314	12.706	31.821	63.657	318.310	636.620
2	0.817	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	0.765	1.638	2.353	3.183	4.541	5.841	10.213	12.924
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.711	1.415	1.895	2.365	2.998	3.500	4.785	5.408
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.700	1.372	1.813	2.228	2.764	3.169	4.144	4.587
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.691	1.341	1.753	2.132	2.602	2.947	3.733	4.073
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.681	1.303	1.684	2.021	2.423	2.705	3.307	3.551
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.677	1.289	1.656	1.980	2.358	2.617	3.160	3.373
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

$$\bar{x}_{min} = \bar{x}' - (t\sigma'/\sqrt{n})$$
$$\bar{x}_{max} = \bar{x}' + (t\sigma'/\sqrt{n})$$

The estimated mean and estimated standard deviation are computed from the equations in 6.4.1 and 6.4.2, and n is the total number of readings of the sample. The value of t is tabulated in Table 2 as a function of the degrees of freedom and the desired degree of confidence. For repeated readings of a given input, the degrees of freedom equals $n-1$ since only one constant, $\bar{x}'$, is estimated from the data. Note that as the number of readings increases, the interval becomes tighter around the estimated mean (by virtue of the factor $1/\sqrt{n}$ and of the value of t), indicating a higher degree of confidence about the true value.

6.5 Chauvenet's Criterion for Rejecting Data Points. See Example 7 in the appendix.

6.5.1 It is a rare circumstance when no data points appear out of place in comparison with the bulk of the data. Therefore, one must decide if these points that appear to be out of place are the result of some gross experimental blunder

and hence can be neglected or if they represent some new type of physical phenomenon that is peculiar to a certain operating condition. The engineer cannot just dismiss points that do not fit his expectations; he must have some consistent basis for elimination.

One such test is known as Chauvenet's criterion. It provides a consistent method of discarding dubious data points. Other criteria for rejecting suspicious data points are reported by Ku.[2]

6.5.2 According to Chauvenet's criterion, a reading may be rejected if the probability of obtaining that particular deviation from the mean is less than $0.5n$. To apply the criterion, a trial mean and trial standard deviation are computed, using all data points. Then the deviations of the individual points are divided by the trial standard deviation and compared with the values given in Table 3. All points whose ratio of deviation to trial standard deviation is higher than that given in Table 3 are rejected, and a new mean and standard deviation are then computed with the rejected points eliminated.

TABLE 3 Chauvenet's Criterion for Rejecting a Reading

Number of Readings, *n*	Ratio of Maximum Acceptable Deviation to Standard Deviation, $dmax/\sigma$
2	1.15
3	1.38
4	1.54
5	1.65
6	1.73
7	1.80
10	1.96
15	2.13
25	2.33
50	2.57
100	2.81
300	3.14
500	3.29
1,000	3.48

6.5.3 Chauvenet's criterion may not be applied a second time to the data set from which points have already been rejected.

6.6 Statistical Tests of a Given Hypothesis

6.6.1 Engineers are frequently required to determine whether the difference between two sets of data is real and significant or if it is due to chance errors. Stated differently, they must determine what the probability is that the two sets of data are drawn from different populations.

Two important statistical tests may be used to answer this question. The t-test is used for small-sample, continuous data (such as temperature, pressure, voltage, etc.) that give a smooth and continuous distribution curve. The Chi-squared test, on the other hand, is designed for frequency or count type of data where the information is in the form of distinct integers (such as number of failures per lot, number of errors per group, etc.).

6.6.2 ***t*-Test (For Fewer Than 20 Data Points).** See Example 8 in the appendix.

In using the *t*-test, we first compare the *t*-value:

where

$$t = |\bar{x}' - \bar{x}|\bar{\sigma}'$$

$\bar{x}'$ = the estimated mean

$\bar{x}$ = the true mean or assumed mean

$\bar{\sigma}' = \sigma'/\sqrt{n}$ = estimated standard deviation of the mean, derived from pooled data.

Table 2 gives values of *t* for various degrees of freedom. The probability level is the chance of getting a *t*-value larger than the table values by chance alone.

In general, the following statements can be made regarding the probability levels for evaluating hypotheses:

(a). For probability levels of 10%, 20%, or 30%, a hypothesis is reasonable (not proved, however).

(b). A 5% probability level raises doubts as to the validity of the hypothesis.

(c). For probability levels less than 1%, a hypothesis can be rejected with confidence.

6.6.3 Chi-Squared Test (Use for 20 or More Data Points). See Example 9 in the appendix.

In general, the chi-squared test tells us whether an observed count differs significantly from an expected count. Chi-squared is defined as

$$\chi^2 = (O - E)^2/E$$

where

O = observed count,

E = expected count.

Table 4 gives chi-squared values for various probability levels and degrees of freedom. Application of the chi-squared test is similar to that of the *t*-test.

7. TREATMENT OF SINGLE-SAMPLE DATA

See Example 3 in Appendix B.

7.1 When the available data are single-sample, the statistical methods of Section 6 are not directly applicable.

7.2 Based on the analysis of Kline and McClintock,[1] a single-sample result should be expressed in terms of a mean value and an uncertainty interval based on a stated confidence level. This result may be written as follows:

$$v = v_m \pm w, (b \text{ percent})$$

where

v_m = the measured value;

w = the uncertainty, computed according to 5.4.4; and

b = the probability that the true value lies within the stated range, based on the opinion of the experimenter.

This approach is of particular value in setting up an experiment, especially if the test is expensive in terms of manpower, time, or equipment. It provides a basis for establishing predetermined estimates of the reliability of results through a study of propagation of uncertainties.

8. DATA PLOTTING AND CURVE FITTING

8.1 Fitting Linear Equations by Least Squares

8.1.1 To increase physical insight into the meaning of empirical results or to present them in a coherent manner, it is customary to plot the data on a graph and to fit a curve to the data by the method of least squares. In order to quantify the scatter of the data and to determine whether any trends exist, it is necessary to obtain a measure of the systematic and random errors that may have been associated with the measurement aspects of the experiment. For the systematic error, the physics of the experiment or instrumentation, or both, must be understood for a complete correction or explanation. Statistics can only suggest that such an error may exist but not why. The random error, which by definition cannot be explained, is properly quantified, for two variables known to be interdependent, by the standard error of estimate.

Figure 4 illustrates a set of data plotted in sequence where the two statistical parameters, mean and standard deviation, as defined in Section 6, are shown. The standard deviation, σ_y, is a measure of the scatter about the mean, $\bar{y}$. When the mean value, $\bar{y}$, for instance, is calculated, in effect a constant has been fitted to the n data points by the method of least squares. That is, a value $\hat{m}$ has been selected for the mean m such that

$$\sum_{1}^{n}(y_i - \hat{m})^2 = \sum_{1}^{n} d_i^2$$

is a minimum. The solution is $\hat{m} = \bar{y}$ and the deviations $d_i = y_i - \hat{m} = y_i - \bar{y}$ are the residuals. The prediction of a value of y that is not one of the data points would then be expressed as $y = M + \varepsilon$ where y is the desired value, m the limiting mean

TABLE 4 Chi-Squared Values

F, Degrees of Freedom	*P*, Probability of Obtaining a Given Value of Chi-Squared (or a larger one)												
	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.001
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	0.102	0.455	1.32	2.671	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.2	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

(estimated by y), and ε the random error with its own limiting mean of zero and a standard deviation, σ_y. The method of least squares requires that the sum of the squares of the residuals divided by the number of measurements, n, less the number of constants, p, in the mathematical equation will define σ_2.

$$\text{Thus: } \sigma^2 = \sum_{i=1}^{n} (y_i - m)^2/(n-p)$$

For Figure 4 this is:

$$\sigma_y^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2/(n-1)$$

which is also the standard deviation defined in Section 6.

Data plotting of this sort (one dimensional) offers little opportunity to increase physical insight into the experiment. It is illustrated here to demonstrate that standard deviation is a special case of one-dimensional least squares fit about a horizontal straight line, while the least squares fit to a sloping straight line (8.1.2) results in what is known as standard error of estimate. This distinction should be noted particularly when using the prepackaged formulae of calculator software or microcomputers.

8.1.2 If a set of experimental results plotted on a graph, such as Figure 5, exhibits an obvious linear dependency of one variable, y, to the independent variable, x, the mathematical expression may be written:

$$y = a + bx + \varepsilon$$

where a (the intercept) and b (the slope) are two constants to be estimated (i.e., $p = 2$) and ε is the random error with a limiting mean of zero and a variance of σ_2. Here again, in accordance with the method of least squares requirements to minimize the sum of the residuals squared, the variance is:

$$\sigma_y^2 = \left[\sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right]/(n-2)$$

where the degrees of freedom, n (i.e., number of data points), are reduced by 2 because they are needed to determine the coefficients a and b. Note that $\hat{y} \neq \bar{y}$, as in 8.1.1, since the functional relationship is not a constant (horizontal line) but rather

$$\hat{y}_i = \hat{a} + \hat{b}x_i$$

where the sum of the squares of the residuals

$$\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

will be a minimum when

$$\hat{b} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) / \sum_{i=1}^{n} (x_i - \bar{x})^2$$

and $\hat{a} = \hat{y} - \hat{b}\bar{x}$.

Thus given a number of data points, it is possible to fit a straight line and to obtain a measure of the scatter, as well as the goodness of fit, by the parameter, σ_y, which is known as the standard error of estimate.

For purposes of obtaining a measure of confidence for the fitted values of the intercept "$\hat{a}$" and the slope "$\hat{b}$," the standard errors of $\hat{b}$ and $\hat{a}$ are used:

$$\sigma_{\hat{b}}^2 = \sigma_y^2/(x_i - x)^2$$
$$\sigma_{\hat{a}}^2 = \sigma_y^2[1/n + \bar{x}^2/\Sigma(x_i - \bar{x})^2]$$

These estimates (i.e., $\sigma_{\hat{b}}^2$ and $\sigma_{\hat{a}}^2$) along with the degrees of freedom associated with σ_y^2 can be used to determine the upper and lower bounds of the coefficients. These bounds are, respectively:

$$\hat{a} \pm t\sigma_{\hat{a}}$$
$$\hat{b} \pm t\sigma_{\hat{b}}$$

where t is obtained from Table 2 for the desired confidence level and the assumed normal distribution curve as illustrated in Figure 5.

8.2 Nonlinear Curve Fitting. See Example 11 in the appendix.

8.2.1 Often a plot of empirical data will exhibit a nonlinear dependency of one variable upon another. In such cases, it is advisable to transform the mathematical expression selected to best fit the data to a linear format. Some common examples of this type of transformation are:

1. An exponential curve: $y = a \cdot e^{bx}$
   $\ln(y) = \ln(a) + bx$
   $y' = a' + bx$
2. A logarithmic curve: $y = a + b\ln(x)$
   $y = a + bx'$

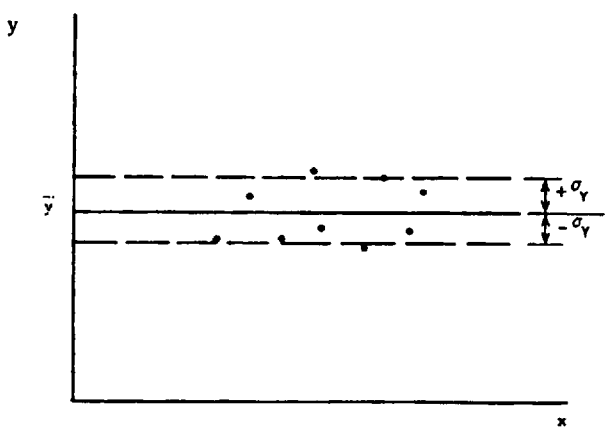

Figure 4 Data plot illustrating one-dimensional statistical parameters.

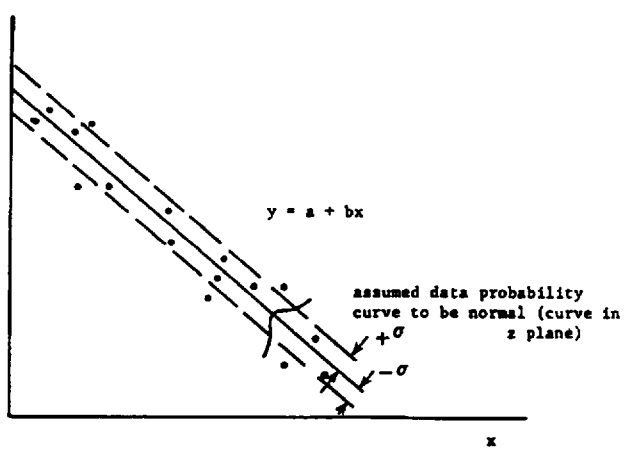

Figure 5 Linear curve fit.

3. A power curve: $y = ax^b$

$$1n(y) = 1n(a) + b\ln(x)$$

$$y' = a' + bx'$$

Polynomial expressions can also be transformed but not without some complexity. Texts on linear regression analysis should be consulted for this.

After the transform is effected, the analysis described in 6.1.2 is performed on the transformed equation. When the variance used to quantify the scatter is determined and the inverse transformed back to the original graph, i.e.,

$$\hat{y} + \varepsilon_2 = e^{(\hat{y}' + \sigma_y)}$$

and

$$\hat{y} - \varepsilon_1 = e^{(\hat{y}' - \sigma_y)}$$

is made, it should not be surprising that the bounds for a given confidence level are neither equal on either side of the fitted curve nor are they constant with the independent variable. This is illustrated in general in Figure 6.

8.2.2 The selection of the proper equation form (i.e., mathematical expression) to be used for the curve fit is first dependent on whether it is known that the one variable is a function of the other or not. In most engineering data analysis, it is known that one variable has a functional dependency on the other. In these cases, the physics of the experiment should suggest the expected proper form (i.e., exponential, power, etc.) prior to fitting data. Of course, the data are needed to determine the specific coefficient values and thus provide the $\hat{y} = f(x)$ quantification necessary to fit the curve to the data points. In theory, it is possible to fit different equation forms to the same data (all transformed to linear format) and select the one with the smallest variance. However, this procedure is not recommended unless it is definitely known that some functional relationship between the plotted parameters does exist and that the independent variable measurement errors are significantly less than the dependent variable. Curve fitting of data plots where both variables have random errors of the same magnitude requires additional analysis, and Chapter 5 of Reference 3 should be consulted. If a purely statistical correlation (i.e., it is not known if a functional relationship exists) is being attempted, the customary unit of measure for deciding whether the data correlate linearly or not is the correlation coefficient, P, or its square, the coefficient of determination.

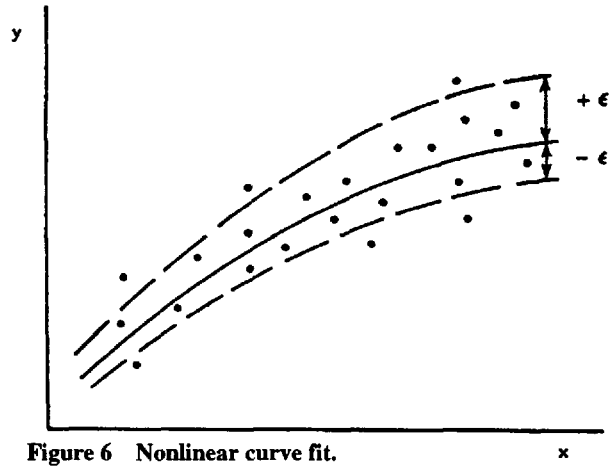

Figure 6 Nonlinear curve fit.

Although it is often stated that a value of P near 1 or −1 suggests good correlation and a value near zero as no correlation, this is an oversimplification and not true in general. The significance of the value is strongly dependent on the number of data points and the level of confidence desired. Futhermore, no degree of statistical correlation can prove functional dependency; that can only come from the physics of the experiment. Before attempting statistical correlations, a thorough familiarity with Chapter 5 of Reference 7 and Reference 8 should be attained.

9. REFERENCES

[1]Kline, S.J., and McClintock, F.A.: Describing uncertainties in single sample experiments, *Mechanical Engineering*, January 1953.

[2]Ku, H.H., ed.: "Statistical Concepts and Procedures," Precision Measurement and Calibration, Vol. 1, Special Publication 300, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C.

[3]Tuve, G.L., and Domholdt, L.C.: *Engineering Experimentation*, McGraw-Hill Book Co., New York, 1966.

[4]Schenck, Jr., H.: *Theories of Engineering Experimentation*, McGraw-Hill Book Co., New York, 1966.

[5]Holman, J.P.: *Experimental Methods for Engineers*, 2d ed., McGraw-Hill Book Co., New York, 1971.

[6]Ku, H.H.: "Statistical Concept in Metrology," Chapter 2, *Handbook of Industrial Metrology*, American Society of Tool Manufacturing Engineers, Prentice-Hall, 1967. (Reprinted as paper 5.2 in Reference 2 cited above.)

[7]Natrella, M.G.: *Experimental Statistics*, *NBS Handbook* 91, U.S. Government Printing Office, Washington, D.C., 1963.

[8]Hahn, G.J.: The coefficient of determination exposed, *Chemical Technology*, Vol. 3, p. 609, American Chemical Society, October 1973.

[9]Barnett, V., and Lewis, T.: *Outlines in Statistical Data*, John Wiley and Sons, New York, 1978.

This Appendix is not part of this Standard but is included for information purposes only.

APPENDIX: EXAMPLE SECTION

Example 1. A reference bath is known to be completely stabilized at a temperature of 100°F (37.8°C). Thermometer A measures the temperature of this bath obtaining in successive measurements 105°F (40.6°C), 95°F (35°C), 100°F (37.8°C), 103°F (39.4°C), 99°F (37.2°C), and 98°F (36.7°C). Thermometer B gives readings of 95.0°F (35.0°C), 95.2°F (35.1°C), 94.8°F (34.9 °C), 95.1°F (35.1°C), and 95.0°F (35.0°C). These results are shown on the figure below. Comment on the accuracy and precision of each.

Thermometer A gives data with relatively large random (precision) error. For this thermometer, the average of the six readings is 100°F (37.8°C). Since this checks the known temperature of 100°F (37.8°C), we infer that little or no systematic (accuracy) error exists.

Thermometer B, on the other hand, has little random error but relatively large systematic error.

In many cases, the experimenter would take a single reading. For Thermometer A, the chance that such a single reading will be accurate is very small indeed. In general, we should replace Thermometer A, whereas Thermometer B could be calibrated to correct the systematic error.

Example 2. Determine the uncertainty in the electric power, P, when determined from the expression:

$$P = EI$$

where

E and I are measured as

E = 100 volts ±3 volts (3% uncertainty)

I = 10 amperes ±0.2 amperes (2% uncertainty)

The nominal value for the power is $100 \times 10 = 1000$ watts.

Based on the method of 5.2.1, the uncertainty in the power would be 3% (i.e., maximum uncertainty in any parameter).

Based on the method of 5.2.2:

$P_{max} = (100 + 3)(10 + 0.2) = 1050.6$ watts

$P_{min} = (100 - 3)(10 - 0.2) = 950.6$ watts

Using this method, the uncertainty is + 5.06%, −4.94%. It is quite unlikely that the power would be in error by these amounts because the maximum voltmeter variation would probably not occur at the same instant as the maximum ammeter variation.

Based on the method of 5.4.4 (recommended):

$$w_p = [(\partial P/\partial E w_E)^2 + (\partial P/\partial I w_1)^2]^{0.5}$$
$$= [(I w_e)^2 + (E w_I)^2]^{0.5}$$

Dividing by $P = EI$:

$$w_p/P = [(w_E)^2/E + (w_I/I)^2]^{0.5}$$
$$= [(0.03)^2 + (0.02)^2]^{0.5} = 0.036 \text{ or } 3.6\%$$

Example 3. An orifice flowmeter is used to measure the flow of air. The expression for obtaining the flow rate is

$$\dot{m} = CA\,[(2g_c p_1/RT_1)\,\Delta p]^{0.5} \qquad (1)$$

where C is an empirical discharge coefficient, A is the flow area, p_1 is the upstream pressure, Δp is the pressure drop across the orifice, T_1 is the upstream temperature, and R is the gas constant for air. Calculate the percent uncertainty in the mass flow rate for the following conditions:

C = 0.92 ± 0.005 (from calibration data)

p_1 = 25 psia ± 0.5 psia

$T_1 = 70°F \pm 2°F$ or $T_1 = 530°R \pm 2°R$

$\Delta_p = p_1 - p_2 = 1.4$ psia ± 0.005 psia (measured directly)

A = 1.0 in.^{2} ± 0.001 in.^{2}

The flow rate is a function of several variables, each subject to an uncertainty:

$$\dot{m} = f(C, A, p_1, \Delta p, T_1)$$

Strictly speaking, p_1, and Δp are not independent. However, because we measure Δp directly, and if we assume the change of Δp with respect to p_1, is quite small, we will assume that they are approximately independent in this case.

Thus, we form the derivatives

$$\partial \dot{m}/\partial c = A[(2g_c p_1/RT_1)\Delta p]^{0.5}$$
$$\partial \dot{m}/\partial A = C[(2g_c p_1/RT_1)\Delta p]^{0.5}$$
$$\partial \dot{m}/\partial p_1 = 0.5\,CA[(2g_c/RT_1)\Delta p]^{0.5}\,p_1^{-0.5}$$
$$\partial \dot{m}/\partial_{\Delta p} = 0.5\,CA[(2g_c p_1/RT_1]^{0.5}\,\Delta p^{-0.5}$$
$$\partial \dot{m}/\partial T_1 = -0.5\,CA(2g_c p_1\,\Delta p\,R)^{0.5}\,T_1^{-1.5}$$

The uncertainty in the mass flow rate may now be calculated by assembling these derivatives in accordance with 5.4.4. Dividing the assembly by Equation 1 yields

$$W_{\dot{m}}/\dot{m} = [(w_c/c)^2 + (w_a/A)^2 + 0.25\,(w_{p1}/p_1)^2 + 0.25\,(w_{\Delta p}/\Delta p)^2 + 0.25\,(w_{T_1}/T_1)^2]^{0.5} \qquad (2)$$

$$W_{\dot{m}}/\dot{m} = [(0.005/0.92)^2 + (0.001/1.0)^2 + 0.25\,(0.5/25)^2 + 0.25\,(0.005/1.4)^2 + 0.25\,(2/530)^2]^{0.5}$$
$$= [29.5 \times 10^{-6} + 1.0 \times 10^{-6} + 1.0 \times 10^{-4} + 3.19 \times 10^{-6} + 3.57 \times 10^{-6}]^{0.5}$$
$$= [1.373 \times 10^{-4}]^{0.5} = 1.172\%$$

The main contribution to uncertainty is the p_1 measurement with its basic uncertainty of 2%. Thus, to improve the overall situation, the accuracy of this measurement should be attacked first.

In order of influence on the flow rate uncertainty, we have

1. Uncertainty in p_1 measurement (±2/%)
2. Uncertainty in value of C
3. Uncertainty in determination of T_1
4. Uncertainty in determination of Δp
5. Uncertainty in determination of A

By inspecting Equation 2, we see that the first two items make practically the whole contribution to uncertainty. The value of the uncertainty analysis in this example is that it shows the investigator how to improve the overall measurement accuracy of the technique. First, one should obtain a more precise measurement of p_1. Then one should try to obtain a better calibration of the device, i.e., a better value of C.

Example 4. The following data were collected for a certain measurement from a Bourdon-tube pressure gage. Plot the data on probability paper and comment on the normality of the distribution.

Reading Number	Gage Reading, psig
1	46.0
2	47.0
3	48.6
4	46.5
5	46.3
6	46.2
7	46.9
8	45.3
9	45.9
10	45.8
	$\Sigma x_i = 464.5$

The mean value is calculated from Section 6.2.3 as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(464.5) = 46.45 \text{ psig}$$

The data are plotted in the figure below and indicate a reasonably normal distribution, although the straight line crosses the 50% ordinate at a value of approximately 46.2 psig, which does differ somewhat from the calculated mean value of 46.45 psig. Comparison with Figure 2 indicates that the distribution is slightly skewed.

Example 5. The mean of 100 measurements is 52.00; the standard deviation of those measurements is 4.00. Find the 90% confidence limits of the population average.

With a sample this large or larger, the estimate of the standard deviation from the sample may be used without correction, and the standard deviation of the mean of the estimate will be $\sigma/\sqrt{n}$, where σ is the estimate of the standard deviation calculated from the sample and n is the number of measurements. In this case the standard deviation of the mean is 4/10 = 0.40.

Using the normal distribution, the table of normal deviates reveals that ± 1.64 standard deviation from the average will include 90% of the values, i.e., Z equals 1.64.

We should then expect that 90% of the time we would be correct in stating that the value of the parameter is within 51.34 to 52.66 (52.00–[1.64 × 0.40] to 52.00 + [1.64 × 0.40]). A less cau-

tious estimate (i.e., 80% confidence interval) would have a narrower confidence interval (Z would be smaller), and a more confident estimate would have larger interval limits.

Example 6. The following thermocouple readings were taken of a stabilized reference bath with various apparatus and procedures. Compute the mean reading, standard deviation, variance, mean deviation (probable error), standard deviation of the mean, and estimate the uncertainty in the calculated best value of the readings.

Reading Number	EMF, mv
1	5.30
2	5.73
3	6.77
4	5.26
5	4.33
6	5.45
7	6.09
8	5.64
9	5.81
10	5.75
11	5.42
12	5.31
13	5.86
14	5.70
15	4.91
16	6.02
17	6.25
18	4.99
19	5.61
20	5.81
21	5.60
$n = 21$	$\Sigma x_i = 117.61$

The mean value is:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{21}(117.61) = 5.60$$

The following table aids in computing the other quantities:

Reading	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	–0.30	0.0900
2	0.13	0.0169
3	1.17	1.3689
4	–0.34	0.1156
5	–1.27	1.6129
6	–0.15	0.0225
7	0.49	0.2401
8	0.04	0.0016
9	0.21	0.0441
10	0.15	0.0225
11	–0.18	0.0324
12	–0.29	0.0841
13	0.26	0.0676
14	0.10	0.0100
15	–0.69	0.4761
16	0.42	0.1764
17	0.65	0.4225
18	–0.61	0.3721
19	0.01	0.0001
20	0.21	0.0441
21	0.0	0
$n = 21$	$\Sigma\lvert x_i - \bar{x}\rvert = 7.67$	$\Sigma(x_i - \bar{x})^2 = 5.2205$

The standard deviation is

$$\sigma = \left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{0.5} = \left[\frac{1}{21}(5.2205)\right]^{0.5} = 0.50 \text{ mv}$$

The average of the readings is 5.60 mv and the standard deviation is 0.50 mv. This means that if we arbitrarily select any single reading, there is (a) a 68.3% chance that it will fall in the region between 6.1 and 5.10 mv ($\bar{x} \pm 1\sigma$); (b) a 95.5% chance that it will fall between 6.60 and 4.60 mv ($\bar{x} \pm 2\sigma$); and (c) a 99.7% chance that it will fall between 7.10 and 4.10 mv ($\bar{x} \pm 3\sigma$).

The variance is:

$$\sigma^2 = 0.25 \text{ mv}$$

The mean deviation ("probable error") is:

$$\bar{d}_i = \frac{1}{n}\sum_{i=1}^{n}|d_i| = \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{1}{21}(7.67) = 0.36 \text{ mv}$$

The standard deviation of the mean is:

$$\bar{\sigma} = \sigma/\sqrt{n} = 0.50/\sqrt{21} = 0.109 \text{ mv}$$

The arithmetic mean (best) value was 5.60 mv with the final result reported in terms of the uncertainty as:

$\bar{x}$ = 5.60 + 0.109 mv; 68.3% (or 2.15 to 1)
= 5.60 + 0.218 mv; 95% (or 21 to 1) (Recommended)
= 5.60 + 0.327 mv; 99.7% (or 256 to 1)

Example 7. Using Chauvenet's criterion, test the data of Example 6 for possible inconsistency. Eliminate questionable points and calculate a new standard deviation.

First, the ratio d_i/σ is calculated and data points eliminated in accordance with Table 3.

Reading	d_i/σ
1	0.60
2	0.26
3	2.34
4	0.68
5	2.54
6	0.30
7	0.98
8	0.08
9	0.42
10	0.30
11	0.36
12	0.58
13	0.52
14	0.20
15	1.38
16	0.84
17	1.30
18	1.22
19	0.02
20	0.42
21	0

In accordance with Table 3 for 21 readings, readings with values of

$d_i/\sigma > 2.2$ are questionable. Thus, only readings 3 and 5 may be eliminated.

Elimination of these points yields a new estimated mean value.

$$x' = \frac{1}{n}\sum_{i=1}^{n} x_i = \left(\frac{1}{19}\right)(106.51) = 5.61$$

Note the change in terminology from mean value to estimated mean value because the sample size has fallen below that required for multisample treatment.

Now the best estimate for the standard deviation is obtained by:

$$\sigma' = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - x')^2\right]^{0.5}$$

The value for σ' is now calculated with the following table:

Reading	$x_i - x'$	$(x_i - x')^2$
1	–0.31	0.0961
2	0.12	0.0144
4	–0.35	0.1225
6	–0.16	0.0256
7	0.48	0.2304
8	0.03	0.0009
9	0.20	0.0400
10	0.14	0.0196
11	–0.19	0.0361
12	–0.30	0.0900
13	0.25	0.0625
14	0.09	0.0081
15	–0.70	0.4900
16	0.41	0.1681
17	0.64	0.4096
18	–0.62	0.3844
19	0.00	0.0000
20	0.20	0.0400
21	0.01	0.0001
n = 19	$\sum\lvert x_i - \bar{x}'\rvert = 5.20$	$\sum\lvert x_i - \bar{x}'\rvert^2 = 2.2384$

$$\sigma' = [(1/18)\,(2.2384)]^{0.5} = 0.35\text{mv}$$

On the basis of the t-distribution, an estimate of the true value of x is:

$$x = \bar{x}' \pm t\sigma'/\sqrt{n}$$

For the 18 degrees of freedom of this problem and the recommended confidence level of 95%

$$t = 2.101$$

and the estimate of the true value of x becomes

$$x = 5.61 \pm 2.101\,(0.35)/\sqrt{19} = 5.61 \pm 0.17 \text{ mv}$$

with the final result reported as

$$x = 5.61 \pm 0.17 \text{ mv; } 95\%$$

Example 8. Measurements of the power supplied to a compressor motor were made on two different days. A summary of the test results is given below:

Aug. 1 (Test A)	Aug. 3 (Test B)
$n = 8$	$n = 17$
$\bar{x} = 3.791$ kW	$\bar{x} = 4.234$ kW
$\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1.270$	$\sum_{i=1}^{n}(x_i - \bar{x})^2 = 1.055$

The test conditions and method of testing were supposedly the same for the two tests. Do these two sets of data belong to the same population?

In comparing two sets of data, we should use a pooled estimate of the standard deviation of the mean in computing the t-value.

$$\bar{\sigma} = \sigma\sqrt{1/n_A + (1/n_B)}$$

where

$$\sigma = \left[\frac{\Sigma(x_A - \bar{x}_A)^2 + \Sigma(x_B - \bar{x}_B)^2}{n_A + n_B - 2}\right]^{0.5}$$

$$F = n_A + n_B - 2 = 23$$

$$\sigma = 0.318 \text{ kW}$$

$$\bar{\sigma} = 0.318\sqrt{\left(\frac{1}{8}\right) + \left(\frac{1}{17}\right)} = 0.136$$

$$t = \lvert(\bar{x}_A - \bar{x}_B)/\bar{\sigma}\rvert = 3.3$$

For $F = 23$ and $t = 3.3$, Table 2 indicates that this t-value corresponds to a probability level of 0.001 and would occur only about once in 1000 times if there were no significant difference between the two tests. Thus we can be fairly certain that either the test conditions or method of testing were not the same for the two tests.

Example 9. A manufacturer of air-conditioning equipment buys motors from two motor manufacturers. Two-thirds are purchased from manufacturer A, and F_A failures are observed. One-third is purchased from manufacturer B, and F_B failures are observed. Is there a significant difference in failure rates between the two motor manufacturers?

The total number of failures is $F_A + F_B$. Since manufacturer A has supplied two-thirds of the motors, we would expect that they would have two-thirds of the failures. Thus, the chi-squared value is

$$\chi^2 = \frac{[F_A - {}^2/_3(F_A + F_B)]^2}{{}^2/_3(F_A + F_B)} + \frac{[F_B - {}^1/_3(F_A + F_B)]^2}{{}^1/_3(F_A + F_B)}$$

$$F = 1$$

Suppose $F_A = 17$ and $F_B = 13$

$$\chi^2 = {}^9/_{20} + {}^9/_{10} = 1.35$$

From Table 4 we can see that this corresponds to a probability level of about 25%. Thus, it is not unreasonable (although not proved) to say that there is no significant difference in failure rates for the motors from manufacturers A and B.

Now suppose that $F_A = 170$ and $F_B = 130$. In this case we have the same ratio of failure rates but a much larger sample.

$$\chi^2 = 13.5$$

Table 4 shows that this value of chi-squared would occur considerably less than 1% of the time, by chance alone, if there were no significant difference in failure rates between A and B. Thus, we can be almost certain that motors from manufacturer B are inferior

to those from manufacturer A.

Example 10. The Energy Policy and Conservation Act (EPCA) establishes an energy conservation program for consumer products. The Department of Energy (DOE) has developed test procedures for the program, which are prescibed to be not unduly burdensome to conduct. Therefore, sampling provisions have been designed both to maximize the confidence with which test results of units actually tested can be applied to similar units not tested and to minimize the testing burden on manufacturers.

The sampling approach recommended by DOE is described by "one-sided confidence limits." The one-sided method places either an upper limit or a lower limit on the range or interval in which the true mean is likely to be found. For example, when measuring something such as the SEER of an air-conditioning unit, for which consumers would favor higher values, the represented value shall be no higher than both the adjusted sample mean and the actual sample mean. The adjusted sample mean is defined as the value for which there is 90% confidence that the true mean is not less than 95% of the adjusted sample mean.

Consider the case in which a manufacturer tests two units from a line of air conditioners. Unit A gave a SEER value of 8.2, and Unit B a SEER value of 8.6. What SEER would the manufacturer be allowed to represent for this line of air conditioners?
The mean value of the sample is

$$\bar{x}' = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{2}(8.2 + 8.6) = 8.4$$

The standard deviation is

$$\sigma' = \left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}')^2\right]^{0.5}$$

$$\sigma' = \left[\frac{1}{1}[(0.2)^2 + (-0.2)^2]\right]^{0.5} = (0.08)^{0.5} = 0.28$$

The lower confidence limit for the true mean can be estimated from the method of 5.4.3.

$$\bar{x} = \bar{x}' - t\sigma'/\sqrt{n}$$

The value of t can be determined from Table 2.

Degrees of freedom = number of observations – number of different quantities

$$= 2 - 1 = 1$$

$$P = (1 - 0.9)2 = 0.2$$

where 0.90 is the confidence level and the factor of 2 is used to convert our "one-sided confidence limit" to the "two-sided confidence limit" on which Table 2 is based.

From Table 2, $t = 3.078$

$$\bar{x} = 8.4 - \frac{3.078 \cdot 0.28}{\sqrt{2}} = 7.79$$

$$\text{Adj. } \bar{x} = \frac{\bar{x}}{0.95} = 8.2$$

Thus, on the basis of these tests the manufacturer could only represent the SEER of the line of air conditioners as 8.2, and it would probably be advantageous for him to conduct a test on a third unit. Suppose Unit C was tested and gave a SEER value of 8.3.

$\bar{x}' = {}^1/_3(8.2 + 8.6 + 8.3) = 8.37$

$\sigma' = [{}^1/_2(-0.17)^2 + (0.23)^2 + (-0.07)^2]^{0.5} = 0.21$

$t = 1.886$

$x = 8.37 - 1.886 \cdot 0.21/\sqrt{3} = 8.14$

Adj. $\bar{x} = 8.14/0.95 = 8.57$

Now, the manufacturer is allowed to represent the SEER of the line of air conditioners as the mean of the sample, or 8.37.

Example 11. A residential air conditioner capacity decay is measured from the moment the compressor and fan in the outdoor unit are shut off. The unit of measure is the ratio of the temperature change across the indoor coil to the steady-state value, where the steady-state value includes the heat from the indoor fan, which is also shut off with the compressor. It is desired to find the best curve fit for the following data:

Time	% of S.S. Value
0 min	88.2
0.25	70.5
0.50	38.1
0.75	31.4
1.00	20.2
1.25	13.2
1.50	8.4
1.75	5.8
2.00	4.6
2.25	1.8

Since the normalized capacity is clearly a nonlinear function of time and experience suggests that capacity decay follows an exponential curve, a least squares fit to a transformed exponential expression is recommended:

$$y = ae^{bx}$$

or

$$y' = a' + bx$$

where

y' = ln (y) = % of S.S. value,
a' = ln (a) = intercept of y' axis at $x = 0$,
b = slope of linearized transformed equation, and
x = time.

The least squares method will evolve an equation:

$$\hat{y}_i' = \hat{a}' + \hat{b}x_i$$

which will describe the fitted curve if the constants are evaluated as follows:

x_i	y_i	y_i'	$(x_i - \bar{x}) - (y_i' - y')$	$(x_i - \bar{x})^2$
0	88.2	4.480	–1.958	1.266
0.25	70.5	4.251	–1.322	0.766
0.50	38.1	3.640	–0.563	0.391
0.75	31.4	3.447	–0.265	0.141
1.00	20.2	3.006	–0.033	0.016
1.25	13.2	2.580	–0.024	0.016
1.50	8.4	2.128	–0.230	0.141
1.75	5.8	1.758	–0.614	0.391
2.00	4.6	1.526	–1.062	0.766
2.25	1.8	0.588	–2.421	1.266
$\bar{x} = 1.125$	$\bar{b}$	$\bar{y}' = 2.740$	$\Sigma = -8.492$	$\Sigma = 5.156$

$$\hat{b} = \frac{\Sigma(x_i - \bar{x}') - (y_i' - y')}{(x_i - \bar{x})^2} = \frac{-8.492}{5.156} = -1.647$$

$$\hat{a}' = y' - bx = 2.740 - (-1.647)1.125 = 4.593$$

This results in the best fitted equation in the transformed state being:

$$\hat{y}' = 4.593 - 1.647\hat{x}$$

By taking the inverse of the transformed term:

$$a = e^{\hat{a}'} = e^{4.593} = 98.79$$

the expression for fitting the original data is:

$$Y = 98.79\, e^{-1.647X}$$

It is now possible to obtain a measure of data scatter by determining the standard error of estimate, as follows:

x_i	y'_i	$\hat{y}'_i$	$(y'_i - \hat{y}'_i)^2$	$\hat{y}_i - e^{\hat{y}i'}$
0	4.480	4.593	0.0128	98.79
0.25	4.251	4.182	0.0048	65.50
0.50	3.640	3.770	0.0169	43.38
0.75	3.447	3.358	0.0079	28.73
1.00	3.006	2.947	0.0035	19.05
1.25	2.580	2.534	0.0021	12.60
1.50	2.128	2.123	0.0000	8.36
1.75	1.758	1.711	0.0022	5.53
2.00	1.526	1.299	0.0515	3.67
2.25	0.588	0.887	0.0894	2.43
			$\Sigma = 0.1911$	

$$\sigma'_y = \left(\frac{\Sigma(y'_i - \hat{y}')^2}{n-2}\right)^{0.5} = \left(\frac{0.1911}{8}\right)^{0.5} = 0.1545$$

This value is for the linear (transformed) equation. To obtain the bounds for the original (inverse) equation, consider the end points:

$$\left.\begin{aligned} y_{10} + \sigma_y &= e^{(\hat{y}'_{10} + \sigma'_y)} = e^{(0.888 + .1545)} = 2.84 \\ y_{10} + \sigma_y &= e^{(\hat{y}'_{10} + \sigma'')} = e^{(0.888 - .1545)} = 2.08 \end{aligned}\right\} 2.43 \begin{matrix} +0.41 \\ -0.35 \end{matrix}$$

The results of these calculations are illustrated in the figure below.

As a matter of rigor it is possible to determine the bounds on the fitted curve coefficients, a and b, with a 95% level of confidence as follows:

$$\sigma_{\hat{b}} = \left[\frac{\sigma'^2_y}{\Sigma(x_i - x)^2}\right]^{0.5} = \left[\frac{1.545^2}{5.156}\right]^{0.5} = 0.680$$

$$\sigma'_{\hat{a}} = \sigma'^2_y\left(\frac{1}{n} - \frac{\bar{x}^2}{\Sigma(x_i - \bar{x})^2}\right)^{0.5}$$

$$= \left[(0.0239)\left(\frac{1}{10} + \frac{(1.125)^2}{5.1563}\right)\right]^{0.5} = 0.0908$$

$b \pm t\sigma_{\hat{b}} = -1.647 \pm (2.306)0.0680 = -1.647 \pm 0.157$ or $\hat{b}$ is within $-1.490 < \hat{b} < -1.804$

$\hat{a}' \pm t\sigma_{\hat{a}'} = 4.593 \pm (2.306)\ 0.0908 = 4.593 \pm 0.209$

taking the inverse $\hat{a} = e^{(4.593 + 0.209)} = 121.8$

$= e^{(4.593 - 0.209)} = 80.2$

or $\hat{a}$ is within $80.2 < \hat{a} < 121.8$.

What becomes obvious from this final calculation is that once the scatter is more or less set with these data points, little is to be gained (i.e., the bounds within which the coefficients are known to exist) by taking more data, since from Table 2 it can be seen that the t-value can only be reduced by about another 15%. Of course, for this problem the physics suggest that $\hat{a}$ is not greater than 100, which helps considerably.

POLICY STATEMENT DEFINING ASHRAE'S CONCERN FOR THE ENVIRONMENTAL IMPACT OF ITS ACTIVITIES

ASHRAE is concerned with the impact of its members' activities on both the indoor and outdoor environment. ASHRAE's members will strive to minimize any possible deleterious effects on the indoor and outdoor environment of the systems and components in their responsibility while maximizing the beneficial effects these systems provide, consistent with accepted standards and the practical state of the art.

ASHRAE's short-range goal is to ensure that the systems and components within its scope do not impact the indoor and outdoor environment to a greater extent than specified by the standards and guidelines as established by itself and other responsible bodies.

As an ongoing goal, ASHRAE will, through its Standards Committee and extensive technical committee structure, continue to generate up-to-date standards and guidelines where appropriate and adopt, recommend, and promote those new and revised standards developed by other responsible organizations.

Through its *Handbook*, appropriate chapters will contain up-to-date standards and design considerations as the material is systematically revised.

ASHRAE will take the lead with respect to dissemination of environmental information of its primary interest and will seek out and disseminate information from other responsible organizations that is pertinent, as guides to updating standards and guidelines.

The effects of the design and selection of equipment and systems will be considered within the scope of the system's intended use and expected misuse. The disposal of hazardous materials, if any, will also be considered.

ASHRAE's primary concern for environmental impact will be at the site where equipment within ASHRAE's scope operates. However, energy source selection and the possible environmental impact due to the energy source and transportation will be considered where possible. Recommendations concerning energy source selection should be made by its members.