# PART 1
# INSTRUMENTATION

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

# CHAPTER 1
# INSTRUMENT STATICS

**Jerry Lee Hall**
**Department of Mechanical Engineering**
**Iowa State University**
**Ames, Iowa**

**Sriram Sundararajan**
**Department of Mechanical Engineering**
**Iowa State University**
**Ames, Iowa**

**Mahmood Naim**
**Union Carbide Corporation**
**Indianapolis, Indiana**

**1 TERMINOLOGY** **3**
1.1 Transducer Characteristics 3
1.2 Definitions 4

**2 STATIC CALIBRATION** **6**
2.1 Calibration Process 6
2.2 Fitting Equations to Calibration Data 6

**3 STATISTICS IN THE MEASUREMENT PROCESS** **9**
3.1 Unbiased Estimates 9
3.2 Sampling 9
3.3 Types of Errors 10
3.4 Propagation of Error or Uncertainty 10
3.5 Uncertainty Interval 12
3.6 Amount of Data to Take 14
3.7 Goodness of Fit 15
3.8 Probability Density Functions 16
3.9 Determination of Confidence Limits on $\mu$ 17
3.10 Confidence Limits on Regression Lines 18
3.11 Inference and Comparison 22

**REFERENCES** **31**

## 1 TERMINOLOGY

### 1.1 Transducer Characteristics

A *measurement system* extracts information about a measurable quantity from some medium of interest and communicates this measured data to the observer. The measurement of any variable is accomplished by an instrumentation system composed of *transducers*. Each transducer is an energy conversion device and requires energy transfer into the device before the variable of interest can be detected.

The Instrument Society of America (ISA) defines *transducer* as ''a device that provides usable output in response to a specified measurand.'' The *measurand* is ''a physical quantity, property or condition which is measured.'' The *output* is ''the electrical quantity, produced by a transducer, which is a function of the applied measurand.''[1]

It should be made very clear that the act of measurement involves transfer of energy between the measured medium and the measuring system and hence the measured quantity



Copyrighted Material
Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

is disturbed to some extent, making a perfect measurement unrealistic. Therefore, pressure cannot be measured without an accompanying change in volume, force cannot be measured without an accompanying change in length, and voltage cannot be measured without an accompanying flow of charge. Instead measures must be taken to minimize the energy transfer from the source to be measured if the measurement is to be accurate.

There are several categorical characteristics for a transducer (or measurement system). When the measurand maintains a steady value or varies very slowly with time, transducer performance can be described in terms of *static characteristics*. Instruments associated with rapidly varying measurands require additional qualifications termed *dynamic characteristics*. Other performance descriptors include *environmental characteristics* (for situations involving varying environmental operating conditions), *reliability characteristics* (related to the life expectancy of the instrument under various operating conditions), *theoretical characteristics* (describing the ideal behavior of the instrument in terms of mathematical or graphical relationships), and *noise characteristics* (external factors that can contribute to the measurement process such as electromagnetic surroundings, humidity, acoustic and thermal vibrations, etc.). In this chapter, we will describe the considerations associated with evaluating numerical values for the static characteristics of an instrument.

## 1.2 Definitions

The description of a transducer and its role in a measuring system is based on most of the definitions that follow. Further details of these definitions can be found in other works.[2–4]

*Static calibration* is the process of measuring the static characteristics of an instrument. This involves applying a range of known values of static input to the instrument and recording the corresponding outputs. The data obtained are presented in a tabular or graphical form.

*Range* is defined by the upper and lower limits of the measured values that an instrument can measure. Instruments are designed to provide predictable performance and, often, enhanced linearity over the range specified.

*Sensitivity* is defined as the change in the output signal relative to the change in the input signal at an operating point. Sensitivity may be constant over the range of the input signal to a transducer or it can vary. Instruments that have a constant sensitivity are called "linear."

*Resolution* is defined as the smallest change in the input signal that will yield a readable change in the output of the measuring system at its operating point.

*Threshold* of an instrument is the minimum input for which there will be an output. Below this minimum input the instrument will read zero.

*Zero* of an instrument refers to a selected datum. The output of an instrument is adjusted to read zero at a predefined point in the measured range. For example, the output of a Celsius thermometer is zero at the freezing point of water; the output of a pressure gage may be zero at atmospheric pressure.

*Zero drift* is the change in output from its set zero value over a specified period of time. Zero drift occurs due to changes in ambient conditions, changes in electrical conditions, aging of components, or mechanical damage. The error introduced may be significant when a transducer is used for long-term measurement.

*Creep* is a change in output occurring over a specific time period while the measurand is held constant at a value other than zero and all environmental conditions are held constant.

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

*Accuracy* is the maximum amount of difference between a measured variable and its true value. It is usually expressed as a percentage of full-scale output. In the strictest sense, accuracy is never known because the true value is never really known.

*Precision* is the difference between a measured variable and the best estimate (as obtained from the measured variable) of the true value of the measured variable. It is a measure of repeatability. Precise measurements have small dispersion but may have poor accuracy if they are not close to the true value. Figure 1*a* shows the differences between accuracy and precision.

*Linearity* describes the maximum deviation of the output of an instrument from a best-fitting straight line through the calibration data. Most instruments are designed so that the output is a linear function of the input. Linearity is based on the type of straight line fitted to the calibration data. For example, least-squares linearity is referenced to that straight line for which the sum of the squares of the residuals is minimized. The

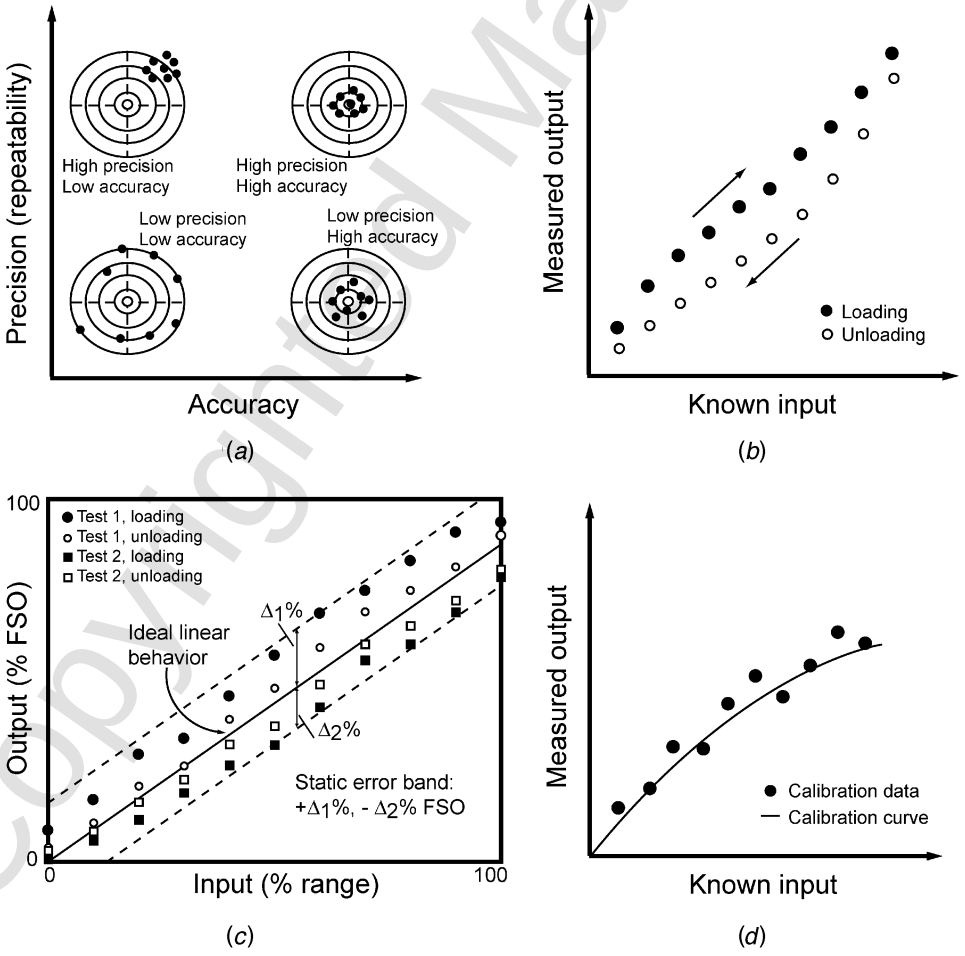



**Figure 1** Schematics illustrating concepts of (*a*) accuracy and precision, (*b*) hysteresis, (*c*) a static error band, and (*d*) fitting a curve to the calibration data.

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

term ''residual'' refers to the deviations of output readings from their corresponding values on the straight line fitted through the data.

*Hysteresis* is the maximum difference in output, at any measured value within the specified range, when the value is approached first with increasing and then with decreasing measurand. Hysteresis is typically caused by a lag in the action of the sensing element of the transducer. Loading the instrument through a cycle of first increasing values, then decreasing values, of the measurand provides a hysteresis loop, as shown in Fig. 1*b*. Hysteresis is usually expressed in percent of full-scale output.

*Error band* is the band of maximum deviation of output values from a specified reference line or curve. A static error band (see Fig. 1*c*) is obtained by static calibration. It is determined on the basis of maximum deviations observed over at least two consecutive calibration cycles so as to include repeatability. Error band accounts for deviations that may be due to nonlinearity, nonrepeatability, hysteresis, zero shift, sensitivity shift, and so forth. It is a convenient way to specify transducer behavior when individual types of deviations need not be specified nor determined.[5]

## 2 STATIC CALIBRATION

### 2.1 Calibration Process

Calibration is the process of comparison of the output of a measuring system to the values of a range of known inputs. For example, a pressure gage is calibrated by a device called a ''dead-weight'' tester, where known pressures are applied to the gage and the output of the gage is recorded over its complete range of operation.

The calibration signal should, as closely as possible, be the same as the type of input signal to be measured. Most calibrations are performed by means of static or level calibration signals since they are usually easy to produce and maintain accurately. However, a measuring system calibrated with static signals may not read correctly when subjected to the dynamic input signals since the natural dynamic characteristics and the response characteristics of the measurement system to the input forcing function would not be accounted for with a static calibration. A measurement system used for dynamic signals should be calibrated using known dynamic inputs.

A static calibration should include both increasing and decreasing values of the known input signal and a repetition of the input signal.[6] This allows one to determine hysteresis as well as the repeatability of the measuring system, as shown in Fig. 1*c*. The sensitivity of the measuring system is obtained from the slope of a suitable line or curve plotted through the calibration points at any level of the input signal.

### 2.2 Fitting Equations to Calibration Data

Though linear in most cases, the calibration plot of a specific measurement system may require a choice of a nonlinear functional form for the relationship that best describes the calibration data, as shown in Fig. 1*d*. This functional form (or curve fit) may be a standard polynomial type or may be one of a transcendental function type. Statistics are used to fit a desired function to the calibration data. A detailed description of the mathematical basis of the selection process used to determine the appropriate function to fit the data can be found elsewhere.[7] Most of today's graphing software allow the user to select the type of fit required. A very common method used to describe the quality of ''fit'' of a chosen functional form is

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

the "least-squares fit." The principle used in making this type of curve fit is to minimize the sum of the squares of the deviations of the data from the assumed curve. These deviations from the assumed curve may be due to errors in one or more variables. If the error is in one variable, the technique is called linear regression and is the common case encountered in engineering measurements. If several variables are involved, it is called multiple regression. Two assumptions are often used with the least-squares method: (i) the $x$ variable (usually the input to the calibration process) has relatively little error as compared to the $y$ (measured) variable and (ii) the magnitude of the uncertainty in $y$ is not dependent on the magnitude of the $x$ variable. The methodology for evaluating calibration curves in systems where the magnitude of the uncertainty in the measured value varies with the value of the input variable can be found elsewhere.[8]

Although almost all graphing software packages include the least-squares fit analysis, thus enabling the user to identify the best-fit curve with minimum effort, a brief description of the mathematical process is given here. To illustrate the least-squares technique, assume that an equation of the following polynomial form will fit a given set of data:

$$y = a + bx^1 + cx^2 + \cdots + mx^k \tag{1}$$

If the data points are denoted by $(x_i, y_i)$, where $i$ ranges from 1 to $n$, then the expression for summation of the residuals is

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

The least-squares method requires that $R$ be minimized. The parameters used for the minimization are the unknown coefficients $a, b, c, \ldots, m$ in the assumed equation. The following differentiation yields $k + 1$ equations called "normal equations" to determine the $k + 1$ coefficients in the assumed relation. The coefficients $a, b, c, \ldots, m$ are found by solving the normal equations simultaneously:

$$\frac{\partial R}{\partial a} = \frac{\partial R}{\partial b} = \frac{\partial R}{\partial c} = \cdots = \frac{\partial R}{\partial m} = 0 \tag{3}$$

For example, if $k = 1$, then the polynomial is of first degree (a straight line) and the normal equations become

$$\Sigma y_i = a(n) + b\Sigma x_i \qquad \Sigma x_i y_i = a\Sigma x_i + b\Sigma x_i^2 \tag{4}$$

and the coefficients $a$ and $b$ are

$$a = \frac{\Sigma x^2 \Sigma y - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} \qquad b = \frac{n\Sigma xy - \Sigma x \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} \tag{5}$$

The resulting curve ($y = a + bx$) is called the regression curve of $y$ on $x$. It can be shown that a regression curve fit by the least-squares method passes through the centroid $(\overline{x}, \overline{y})$ of the data.[9] If two new variables $X$ and $Y$ are defined as

$$X = x - \overline{x} \qquad \text{and} \qquad Y = y - \overline{y} \tag{6}$$

then

$$\Sigma X = 0 = \Sigma Y \tag{7}$$

Substitution of these new variables in the normal equations for a straight line yields the following result for $a$ and $b$:

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$a = 0 \qquad b = \frac{\sum XY}{\sum X^2} \tag{8}$$

The regression line becomes

$$Y = bX \tag{9}$$

The technique described above will yield a curve based on an assumed form that will fit a set of data. This curve may not be the best one that could be found, but it will be the best based on the assumed form. Therefore, the ''goodness of fit'' must be determined to check that the fitted curve follows the physical data as closely as possible.

**Example 1 Choice of Functional Form.** Find a suitable equation to represent the following calibration data:

$x$ = [3, 4, 5, 7, 9, 12, 13, 14, 17, 20, 23, 25, 34, 38, 42, 45]

$y$ = [5.5, 7.75, 10.6, 13.4, 18.5, 23.6, 26.2, 27.8, 30.5, 33.5, 35, 35.4, 41, 42.1, 44.6, 46.2]

***Solution:*** A computer program can be written or graphing software (e.g., Microsoft Excel) used to fit the data to several assumed forms, as given in Table 1. The data can be plotted and the best-fitting curve selected on the basis of minimum residual error, maximum correlation coefficient, or smallest maximum absolute deviation, as shown in Table 1.

The analysis shows that the assumed equation $y = a + (b)\log(x)$ represents the best fit through the data as it has the smallest maximum deviation and the highest correlation coefficient. Also note that the equation $y = 1/a + bx$ is not appropriate for these data because it has a negative correlation coefficient.

**Example 2 Nonlinear Regression.** Find the regression coefficients $a$, $b$, and $c$ if the assumed behavior of the $(x, y)$ data is $y = a + bx + cx^2$:

$x$ = [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]

$y$ = [0.26, 0.38, 0.55, 0.70, 1.05, 1.36, 1.75, 2.20, 2.70, 3.20, 3.75, 4.40, 5.00, 6.00]

**Table 1** Statistical Analysis for Example 1

| Assumed Equation | Regression Coefficient | | Residual Error, $R$ | Maximum Deviation | Correlation Coefficient[a] |
|---|---|---|---|---|---|
| | $a$ | $b$ | | | |
| $y = bx$ | —- | 1.254 | 56.767 | 10.245 | 0.700 |
| $y = a + bx$ | 9.956 | 0.907 | 20.249 | 7.178 | 0.893 |
| $y = ae^{bx}$ | 10.863 | 0.040 | 70.274 | 18.612 | 0.581 |
| $y = 1/(a + bx)$ | 0.098 | −0.002 | 14257.327 | 341.451 | −74.302 |
| $y = a + b/x$ | 40.615 | −133.324 | 32.275 | 9.326 | 0.830 |
| $y = a + b \log x$ | −14.188 | 15.612 | 1.542 | 2.791 | 0.992 |
| $y = ax^b$ | 3.143 | 0.752 | 20.524 | 8.767 | 0.892 |
| $y = x/(a + bx)$ | 0.496 | 0.005 | 48.553 | 14.600 | 0.744 |

[a]Defined in Section 4.3.7.

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

From Eqs. (1)–(3)

$$y = a + bx + cx^2$$
$$xy = ax + bx^2 + cx^3 \tag{10}$$
$$x^2y = ax^2 + bx^3 + cx^4$$

A simultaneous solution of the above equations provides the desired regression coefficients:

$$a = 0.1959 \qquad b = -0.0205 \qquad c = 0.0266 \tag{11}$$

As was mentioned previously, a measurement process can only give you the best estimate of the measurand. In addition, engineering measurements taken repeatedly under seemingly identical conditions normally show variations in measured values. A statistical treatment of measurement data is therefore a necessity.

## 3 STATISTICS IN THE MEASUREMENT PROCESS

### 3.1 Unbiased Estimates

Data sets typically have two very important characteristics: *central tendency* (or most representative value) and *dispersion* (or scatter). Other characteristics such as skewness and kurtosis (or peakedness) may also be of importance but will not be considered here.[10]

A basic problem in every quantitative experiment is that of obtaining an unbiased estimate of the true value of a quantity as well as an unbiased measure of the dispersion or uncertainty in the measured variable. Philosophically, in any measurement process a deterministic event is observed through a ''foggy'' window. If so, ultimate refinement of the measuring system would result in all values of measurements to be the true value $\mu$. Because errors occur in all measurements, one can never exactly measure the true value of any quantity. Continued refinement of the methods used in any measurement will yield closer and closer approximations, but there is always a limit beyond which refinements cannot be made. To determine the relation that a measured value has with the true value, we must specify the unbiased estimate $\bar{x}$ of the true value $\mu$ of a measurement and its uncertainty (or precision) interval $W_x$ based on a desired confidence level (or probability of occurrence).

An *unbiased estimator*[9] exists if the mean of its distribution is the same as the quantity being estimated. Thus, for sample mean $\bar{x}$ to be an unbiased estimator of population mean $\mu$, the mean of the distribution of sample means, $\bar{\bar{x}}$, must be equal to the population mean.

### 3.2 Sampling

Unbiased estimates for determining population mean, population variance, and variance of the sample mean depend on the type of sampling procedure used.

*Sampling with Replacement* (*Random Sampling*)

$$\hat{\mu} = \bar{x} \tag{12}$$

where $\bar{x}$ is the sample mean and $\hat{\mu}$ is the unbiased estimate of the population mean, $\mu$;

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$\hat{\sigma}^2 = S^2\left(\frac{n}{n-1}\right)$$

where $S^2$ is the sample variance

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n} \tag{13}$$

and

$$\hat{\sigma}_{\bar{x}}^2 = \frac{\hat{\sigma}^2}{n} \tag{14}$$

where $\hat{\sigma}_{\bar{x}}^2$ is the variance of the mean.

*Sampling without Replacement* (*Usual Case*)

$$\hat{\mu} = \bar{x} \tag{15}$$

$$\hat{\sigma}^2 = S^2\left(\frac{n}{n-1}\right)\left(\frac{N-1}{N}\right) \tag{16}$$

where $N$ is the population size and $n$ the sample size, and

$$\hat{\sigma}_{\bar{x}}^2 = \frac{\hat{\sigma}^2}{n}\left(\frac{N-n}{N-1}\right) \tag{17}$$

Note that sampling without replacement from an extremely large population is equivalent to random sampling.

## 3.3 Types of Errors

There are at least three types of errors that one must consider in making measurements. They are systematic (or fixed) errors, illegitimate errors (or mistakes), and random errors:

*Systematic errors* are of consistent form. They result from conditions or procedures that are correctable. This type of error may generally be eliminated by calibration.

*Illegitimate errors* are mistakes and should not exist. They may be eliminated by using care in the experiment, proper measurement procedures, and repetition of the measurement.

*Random errors* are accidental errors that occur in all measurements. They are characterized by their inconsistent nature, and their origin cannot be determined in the measurement process. These errors are estimated by statistical analysis.

If the illegitimate errors can be eliminated by care and proper measurement procedures and the systematic errors can be eliminated by calibrating the measurement system, then the random errors remain to be determined by statistical analysis to yield the precision of the measurement.

## 3.4 Propagation of Error or Uncertainty

In many cases the desired quantity cannot be measured directly but must be calculated from the most representative value (e.g., the mean) of two or more measured quantities. It is desirable to know the uncertainty or precision of such calculated quantities.

Precision is specified by quantities called *precision indexes* (denoted by $W_x$) that are calculated from the random errors of a set of measurements. A $\pm W_x$ should be specified for

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

every measured variable. The confidence limits or probability for obtaining the range $\pm W_x$ is generally specified directly or is implied by the particular type of precision index being used.

The precision index of a calculated quantity depends on the precision indexes of the measured quantities required for the calculations.[9] If the measured quantities are determined independently and if their distribution about a measure of central tendency is approximately symmetrical, the following ''propagation-of-error'' equation is valid[11]:

$$W_R^2 = \Sigma\left(\frac{\partial R}{\partial x_i}\right)^2 W_{x_i}^2 \tag{18}$$

In this equation, $R$ represents the calculated quantity and $x_1, x_2, \ldots, x_n$ represent the measured independent variables so that mathematically we have $R = f(x_1, x_2, \ldots, x_n)$. The precision index is a measure of dispersion about the central tendency and is denoted by $W$ in Eq. (18). The standard deviation is often used for $W$; however, any precision index will do as long as the same type of precision index is used in each term of the equation.

A simplified form of this propagation-of-error equation results if the function $R$ has the form

$$R = kx_1^a x_2^b x_3^c \cdots x_n^m \tag{19}$$

where the exponents $a, b, \ldots, m$ may be positive or negative, integer or noninteger. The simplified result for the precision $W_R$ in $R$ is

$$\left(\frac{W_R}{R}\right)^2 = a^2\left(\frac{W_{x_1}}{x_1}\right)^2 + b^2\left(\frac{W_{x_2}}{x_2}\right)^2 + \cdots + m^2\left(\frac{W_{x_n}}{x_n}\right)^2 \tag{20}$$

The propagation-of-error equation is also used in planning experiments. If a certain precision is desired on the calculated result $R$, the precision of the measured variables can be determined from this equation. Then, the cost of a proposed measurement system can be determined as it is directly related to precision.

**Example 3 Propagation of Uncertainty.** Determine the resistivity and its uncertainty for a conducting wire of circular cross section from the measurements of resistance, length, and diameter. Given

$$R = \rho\frac{L}{A} = \rho\frac{4L}{\pi D^2} \qquad \text{or} \qquad \rho = \frac{\pi D^2 R}{4L} \tag{21}$$

$$R = 0.0959 \pm 0.0001\ \Omega \qquad L = 250 \pm 2.5\ \text{cm} \qquad D = 0.100 \pm 0.001\ \text{cm}$$

where $R$ = wire resistance, $\Omega$
$L$ = wire length, cm
$A$ = cross-sectional area, $= \pi D^2/4$, cm$^2$
$\rho$ = wire resistivity, $\Omega\cdot$cm

***Solution:*** Thus the resistivity is

$$\rho = \frac{(\pi)(0.100)^2(0.0959)}{4(250)} = 3.01 \times 10^{-6}\ \Omega\cdot\text{cm}$$

The propagation of variance (or precision index) equation for $\rho$ reduces to the simplified form, that is,

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$\left(\frac{W_\rho}{\rho}\right)^2 = 4\left(\frac{W_D}{D}\right)^2 + \left(\frac{W_R}{R}\right)^2 + \left(\frac{W_L}{L}\right)^2$$
$$= 4\left(\frac{0.001}{0.10}\right)^2 + \left(\frac{0.0001}{0.0959}\right)^2 + \left(\frac{2.5}{250}\right)^2$$
$$= 4.00 \times 10^{-4} + 1.09 \times 10^{-6} + 1.00 \times 10^{-4}$$
$$= 5.01 \times 10^{-4}$$

The resulting resistivity $\rho$ and its precision $W_\rho$ are

$$W_\rho = \rho\sqrt{(5.01)10^{-4}} = \pm 6.74 \times 10^{-8}$$
$$\rho = (3.01 \pm 0.07) \times 10^{-6}\ \Omega\cdot\text{cm}$$

## 3.5 Uncertainty Interval

When several measurements of a variable have been obtained to form a data set (multisample data), the best estimates of the most representative value (mean) and dispersion (standard deviation) are obtained from the formulas in Section 3.2. When a single measurement exists (or when the data are taken so that they are equivalent to a single measurement), the standard deviation cannot be determined and the data are said to be ''single-sample'' data. Under these conditions the only estimate of the true value is the single measurement, and the uncertainty interval must be estimated by the observer.[12] It is recommended that the precision index be estimated as the maximum reasonable error. This corresponds approximately to the 99% confidence level associated with multisample data.

***Uncertainty Interval Considering Random Error***

Once the unbiased estimates of mean and variance are determined from the data sample, the uncertainty interval for $\mu$ is

$$\mu = \hat{\mu} \pm \hat{W} = \hat{\mu} \pm k(\nu, \gamma)\hat{\sigma} \tag{22}$$

where $\hat{\mu}$ represents the most representative value of $\mu$ from the measured data and $\hat{W}$ is the uncertainty interval or precision index associated with the estimate of $\mu$. The magnitude of the precision index or uncertainty interval depends on the confidence level $\gamma$ (or probability chosen), the amount of data $n$, and the type of probability distribution governing the distribution of measured items.

The uncertainty interval $\hat{W}$ can be replaced by $k\hat{\sigma}$, where $\hat{\sigma}$ is the standard deviation (measure of dispersion) of the population as estimated from the sample and $k$ is a constant that depends on the probability distribution function, the confidence level $\gamma$, and the amount of data $n$. For example, with a Gaussian distribution the 95% confidence limits are $\hat{W} = 1.96\sigma$, where $k = 1.96$ and is independent of $n$. For a $t$-distribution, $k = 2.78$, 2.06, and 1.96 with a sample size of 5, 25, and $\infty$, respectively, at the 95% level of confidence probability. Note that $\nu = n - 1$ for the $t$-distribution. The $t$-distribution is the same as the Gaussian distribution as $n \to \infty$.

***Uncertainty Interval Considering Random Error with Resolution, Truncation, and Significant Digits***

The uncertainty interval $\hat{W}$ in Eq. (22) assumes a set of measured values with only random error present. Furthermore, the set of measured values is assumed to have unbounded significant digits and to have been obtained with a measuring system having infinite resolution. When finite resolution exists and truncation of digits occurs, the uncertainty interval may be larger than that predicted by consideration of the random error only. The uncertainty interval can never be less than the resolution limits or truncation limits of the measured values.

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

***Resolution and Truncation***

Let $\{s_n\}$ be the theoretically possible set of measurements of unbound significant digits from a measuring system of infinite resolution and let $\{x_n\}$ be the actual set of measurements expressed to $m$ significant digits from a measuring system of finite resolution. Then the quantity $s_i - x_i = \pm e_i$ is the resolution or truncation deficiency caused by the measurement process. The unbiased estimates of mean and variance are

$$\hat{\mu} = \frac{\sum s_i}{n} = \bar{s} \qquad \text{and} \qquad \hat{\sigma}^2 = \frac{\sum (s_i - \bar{s})^2}{n - 1} \tag{23}$$

Noting that the set $\{x_n\}$ is available rather than $\{s_n\}$, the required mean and variance are

$$\hat{\mu} = \frac{\sum x_i}{n} \pm \frac{\sum e_i}{n} = \bar{x} \pm \frac{\sum e_i}{n} \qquad \text{and} \qquad \hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \tag{24}$$

The truncation or resolution has no effect on the estimate of variance but does affect the estimate of the mean. The truncation error $e_i$ is not necessarily distributed randomly and may all be of the same sign. Thus $\bar{x}$ can be biased as much as $\sum e_i/n = \bar{e}$ high or low from the unbiased estimate of the value of $\mu$ so that $\hat{\mu} = \bar{x} \pm \bar{e}$.

If $e_i$ is a random variable observed through a "cloudy window" with a measuring system of finite resolution, the value of $e_i$ may be plus or minus but its upper bound is $R$ (the resolution of the measurement). Thus the resolution error is no larger than $R$ and $\hat{\mu} = \bar{x} \pm Rn/n = \bar{x} \pm R$.

If the truncation is never more than that dictated by the resolution limits ($R$) of the measurement system, the uncertainty in $\bar{x}$ as a measure of the most representative value of $\mu$ is never larger than $R$ plus the uncertainty due to the random error. Thus $\hat{\mu} = \bar{x} \pm (\hat{W} + \mathrm{R})$. It should be emphasized that the uncertainty interval can never be less than the resolution bounds of the measurement. The resolution bounds cannot be reduced without changing the measurement system.

***Significant Digits***

When $x_i$ is observed to $m$ significant digits, the uncertainty (except for random error) is never more than $\pm 5/10^m$ and the bounds on $s_i$ are equal to $x_i \pm 5/10^m$ so that

$$x_i - \frac{5}{10^m} < s_i < x_i + \frac{5}{10^m} \tag{25}$$

The relation for $\hat{\mu}$ for $m$ significant digits is then from Eq. 24.

$$\hat{\mu} = \bar{x} \pm \frac{\sum e_i}{n} = \bar{x} \pm \frac{\sum (5/10^m)}{n} = \bar{x} \pm \frac{5}{10^m} \tag{26}$$

The estimated value of variance is not affected by the constant magnitude of $5/10^m$. When the uncertainty due to significant digits is combined with the resolution limits and random error, the uncertainty interval on $\hat{\mu}$ becomes

$$\hat{\mu} = \bar{x} \pm \left( \hat{W} + R + \frac{5}{10^m} \right) \tag{27}$$

This illustrates that the number of significant digits of a measurement should be carefully chosen in relation to the resolution limits of the measuring system so that $5/10^m$ has about the same magnitude as $R$. Additional significant digits would imply more accuracy to the measurement than would actually exist based on the resolving ability of the measuring system.

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

## 3.6 Amount of Data to Take

Exactly what data to take and how much data to take are two important questions to be answered in any experiment. Assuming that the correct variables have been measured, the amount of data to be obtained can be determined by using the relation

$$\mu = \bar{x} \pm \left( \hat{W}_{\bar{x}} + R + \frac{5}{10^m} \right) \tag{28}$$

where it is presumed that several sample sets may exist for estimation of $\mu$ and that the mean of means of the sample sets is denoted by $\bar{\bar{x}}$. This equation can be rewritten using Eqs. (13) and (14) (assuming random sampling):

$$\begin{aligned} \mu &= \bar{\bar{x}} \pm \left( k(\nu,\gamma)\hat{\sigma}_{\bar{x}} + R + \frac{5}{10^m} \right) \\ &= \bar{\bar{x}} \pm \left( k(\nu,\gamma)\frac{\hat{\sigma}}{\sqrt{n}} + R + \frac{5}{10^m} \right) \end{aligned} \tag{29}$$

The value of $n$ to achieve the difference in $\mu - \bar{\bar{x}}$ within a stated percent of $\mu$ can be determined from

$$n = \left[ \frac{k(v,\gamma)\hat{\sigma}}{(\%/100)\hat{\mu} - R - (5/10^m)} \right]^2 \tag{30}$$

This equation can only yield valid values of $n$ once valid estimates of $\hat{\mu}$, $\hat{\sigma}$, $k$, $R$, and $m$ are available. This means that the most correct values of $n$ can only be obtained once the measurement system and data-taking procedure have been specified so that $R$ and $m$ are known. Furthermore, either a preliminary experiment or a portion of the actual experiment should be performed to obtain good estimates of $\hat{\mu}$ and $\hat{\sigma}$. Because $k$ depends not only on the type of distribution the data follows but also on the sample size $n$, the solution is iterative. Thus, the most valid estimates of the amount of data to take can only be obtained after the experiment has begun. However, the equation can be quite useful for prediction purposes if one wishes to estimate values of $\hat{\mu}$, $\hat{\sigma}$, $k$, $R$, and $m$. This is especially important in experiments for which the cost of a single run may be relatively high.

**Example 4 Amount of Data to Take.** The life for a certain type of automotive tire is to be established. The mean and standard deviation of the life estimated for these tires are 84,000 and ±7,230 km, respectively, from a sample of nine tires. On the basis of the sample, how much data are required to establish the life of this type of tire to within ±10% with 90% confidence and a resolution of 5 km?

***Solution:*** Confidence limits are as follows:

$$\mu = \hat{\mu} \pm \left( t\hat{\sigma}_{\bar{x}} + R + \frac{5}{10^m} \right)$$

$$\mu - \hat{\mu} = (0.10)\bar{x} = t\hat{\sigma}_{\bar{x}} + R + \frac{5}{10^m} = t\frac{\hat{\sigma}_x}{\sqrt{n}} + R + \frac{5}{10^m}$$

$$\frac{t}{\sqrt{n}} = \frac{(0.10)\bar{x} - R - (5/10^m)}{\hat{\sigma}_x} = \frac{(0.10)(84000) - 5}{7230} = 1.6$$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

| $n$ | $\nu$ | $t(\nu,0.10)^a$ | $t/\sqrt{n}$ |
|---|---|---|---|
| 2 | 1 | 6.31 | 3.65 |
| 3 | 2 | 2.92 | 1.46 |
| 5 | 4 | 2.13 | 0.87 |

[a]From a *t*-statistic table.[9]

Thus a sample of three tires is sufficient to establish the tire life within $\pm 10\%$ at a 90% level of confidence.

## 3.7 Goodness of Fit

Statistical methods can be used to fit a curve to a given data set. In general, the least-squares principle is used to minimize the sum of the squares of deviations away from the curve to be fitted. The deviations from an assumed curve $y = f(x)$ are due to errors in $y$, in $x$, or in both $y$ and $x$. In most cases the errors in the independent variable $x$ are much smaller than the dependent variable $y$. Therefore, only the errors in $y$ are considered for the least-squares curve. The goodness of fit of an assumed curve is defined by the correlation coefficient $r$, where

$$r = \pm \sqrt{\frac{\sum(y - \bar{y})^2}{\sum(y_i - \bar{y})^2}} = \pm \sqrt{1 - \frac{\sum(y_i - y)^2}{\sum(y_i - \bar{y})^2}}$$

$$= \pm \sqrt{1 - \frac{\hat{\sigma}^2_{y,x}}{\hat{\sigma}^2_y}} \tag{31}$$

where $\sum(y_i - \bar{y})^2$ = total variation (variation about mean)
$\sum(y_i - y)^2$ = unexplained variation (variation about regression)
$\sum(y - \bar{y})^2$ = explained variation (variation based on assumed regression equation)
$\hat{\sigma}_y$ = estimated population standard deviation of $y$ variable
$\hat{\sigma}_{y,x}$ = standard error of estimate of $y$ on $x$

When the correlation coefficient $r$ is zero, the data cannot be explained by the assumed curve. However, when $r$ is close to $\pm 1$, the variation of $y$ with respect to $x$ can be explained by the assumed curve and a good correlation is indicated between the variables $x$ and $y$.

The probabilistic statement for the goodness-of-fit test is given by

$$P[r_{\text{calc}} > r] = \alpha = 1 - \gamma \tag{32}$$

where $r_{\text{calc}}$ is calculated from Eq. (31) and the null and alternate hypotheses are as follows:

$H_0$: No correlation of assumed regression equation with data.

$H_1$: Correlation of regression equation with data.

The goodness-of-fit for a straight line is determined by comparing $r_{\text{calc}}$ with the value of $r$ obtained at $n - 2$ degrees of freedom at a selected confidence level $\gamma$ from tables. If $r_{\text{calc}} > r$, the null hypothesis is rejected and a significant fit of the data within the confidence level specified is inferred. However, if $r_{\text{calc}} < r$, the null hypothesis cannot be rejected and no correlation of the curve fit with the data is inferred at the chosen confidence level.

**Example 5 Goodness-of-Fit Test.** The given $x$–$y$ data were fitted to a curve $y = a + bx$ by the method of least-squares linear regression. Determine the goodness of fit at a 5% significance level (95% confidence level):

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$x = [56, 58, 60, 70, 72, 75, 77, 77, 82, 87, 92, 104, 125]$$

$$y = [51, 60, 60, 52, 70, 65, 49, 60, 63, 61, 64, 84, 75]$$

At $\alpha = 0.05$, the value of $r$ is 0.55. Thus $P[r_{calc} > 0.55] = 0.05$. The least-squares regression equation is calculated to be $y = 39.32 + 0.30x$, and the correlation coefficient is calculated as $r_{calc} = 0.61$. Therefore, a satisfactory fit of the regression line to that data is inferred at the 5% significance level (95% confidence level).

## 3.8 Probability Density Functions

Consider the measurement of a quantity $x$. Let

$x_i$ = $i$ measurement of quantity
$\bar{x}$ = most representative value of measured quantity
$d_i$ = deviation of $i$th measurement from $\bar{x}$, $= x_i - \bar{x}$
$n$ = total number of measurements
$\Delta x$ = smallest measurable change of $x$, also known as "least count"
$m_j$ = number of measurements in $x_j$ size group

A *histogram* and the corresponding *frequency polygon* obtained from the data are shown in Fig. 2, where the $x_j$ size group is taken to be inclusive at the lower limit and exclusive at the upper limit. Thus, each $x_j$ size group corresponds to the following range of values:

$$x_j - \left(\frac{1}{2}\Delta x_j\right) \leq x < x_j + \left(\frac{1}{2}\Delta x_j\right) \tag{33}$$

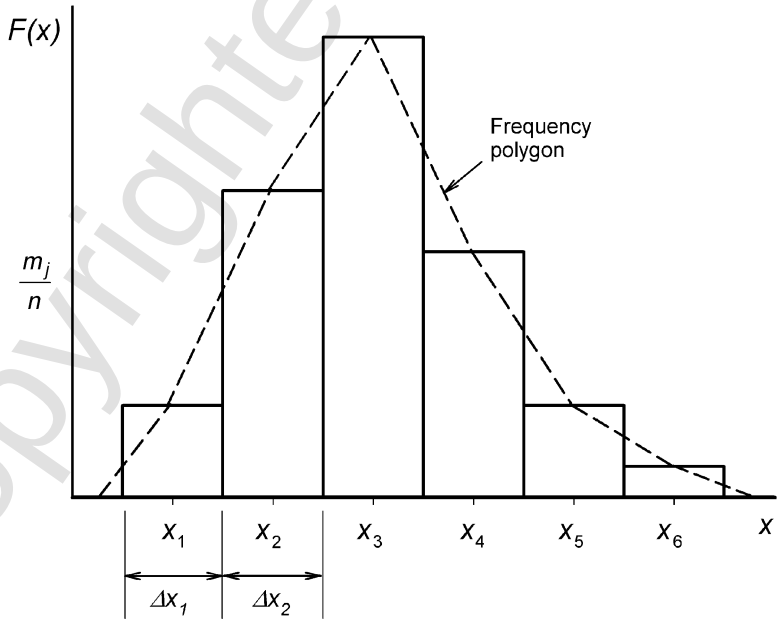



**Figure 2** Histogram of a data set and the corresponding frequency polygon.

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

The height of any rectangle of the histogram shown is denoted as the relative number $m_j/n$ and is equal to the statistical probability $F_j(x)$ that a measured value will have the size $x_j \pm (\Delta x_j/2)$. The area of each rectangle can be made equal to the relative number by transforming the ordinate of the histogram in the following way: Area = relative number = $F_j(x) = m_j/n = p_j(x)\ \Delta x$. Thus

$$p_j(x) = \frac{m_j}{n\ \Delta x} \tag{34}$$

The shape of the histogram is preserved in this transformation since the ordinate is merely changed by a constant scale factor $(1/\Delta x)$. The resulting diagram is called the *probability density diagram.* The sum of the areas underneath all rectangles is then equal to 1 [i.e., $\sum p_j(x)\ \Delta x_j = 1$]. In the limit, as the number of data approaches infinity and the least count becomes very small, a smooth curve called the *probability density function* is obtained. For this smooth curve we note that

$$\int_{-\infty}^{+\infty} p(x)\ dx = 1 \tag{35}$$

and that the probability of any measurement, $x$, having values between $x_a$ and $x_b$ is found from

$$P(x_a \leq x \leq x_b) = \int_{x_a}^{x_b} p(x)\ dx \tag{36}$$

To integrate this expression, the exact probability density function $p(x)$ is required. Based on the assumptions made, several forms of frequency distribution laws have been obtained. The distribution of a proposed set of measurements is usually unknown in advance. However, the Gaussian (or normal) distribution fits observed data distributions in a large number of cases. The Gaussian probability density function is given by the expression

$$p(x) = (1/\sigma\sqrt{2\pi})e^{-(x-\bar{x})^2/2\sigma^2} \tag{37}$$

where $\sigma$ is the standard deviation. The standard deviation is a measure of dispersion and is defined by the relation

$$\sigma = \frac{\int_{-\infty}^{+\infty} x^2 p(x)\ dx}{\int_{-\infty}^{+\infty} p(x)\ dx} = \sqrt{\frac{\sum(x_i - \mu)^2}{n}} \tag{38}$$

## 3.9 Determination of Confidence Limits on $\mu$

If a set of measurements is given by a random variable $x$, then the central limit theorem[13] states that the distribution of means, $\bar{x}$, of the samples of size $n$ is Gaussian (normal) with mean $\mu$ and variance $\sigma_{\bar{x}}^2 = \sigma^2/n$, that is, $\bar{x} \sim G(\mu, \sigma^2/n)$. [Also, the random variable $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$ is Gaussian with a mean of zero and a variance of unity, that is, $z \sim G(0,1)$.] The random variable $z$ is used to determine the confidence limits on $\mu$ due to random error of the measurements when $\sigma$ is known.

The confidence limit is determined from the following probabilistic statement and the Gaussian distribution for a desired confidence level $\gamma$:

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$P\left[-z < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < z\right] = \gamma \tag{39}$$

It shows a $\gamma$ probability or "confidence level" that the experimental value of $z$ will be between $\pm z$ obtained from the Gaussian distribution table. For a 95% confidence level, $z = \pm 1.96$ from the Gaussian table[9] and $P[-1.96 < z < 1.96] = \gamma$, where $z = (\bar{x} - \mu)/(\sigma/\sqrt{n})$. Therefore, the expression for the 95% confidence limits on $\mu$ is

$$\mu = \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} \tag{40}$$

In general

$$\mu = \bar{x} \pm k \frac{\sigma}{\sqrt{n}} \tag{41}$$

where $k$ is found from the Gaussian table for the specified value of confidence, $\gamma$.

If the population variance is not known and must be estimated from a sample, the statistic $(\bar{x} - \mu)/(\sigma/\sqrt{n})$ is not distributed normally but follows the $t$-distribution. When $n$ is very large, the $t$-distribution is the same as the Gaussian distribution. When $n$ is finite, the value of $k$ is the "$t$" value obtained from the $t$-distribution table.[9] The probabilistic statement then becomes

$$P\left[-t < \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < +t\right] = \gamma \tag{42}$$

and the inequality yields the expression for the confidence limits on $\mu$:

$$\mu = \bar{x} \pm t \frac{\sigma}{\sqrt{n}} \tag{43}$$

If the effects of resolution and significant digits are included, the expression becomes as previously indicated in Eq. (29):

$$\mu = \bar{x} \pm \left(t \frac{\sigma}{\sqrt{n}} + R + \frac{5}{10^m}\right) \tag{44}$$

## 3.10 Confidence Limits on Regression Lines

The least-squares method is used to fit a straight line to data that are either linear or transformed to a linear relation.[9,14] The following method assumes that the uncertainty in the variable $x$ is negligible compared to the uncertainty in the variable $y$ and that the uncertainty in the variable $y$ is independent of the magnitude of the variable $x$. Figure 3 and the definitions that follow are used to obtain confidence levels relative to regression lines fitted to experimental data:

$a$ = intercept of regression line
$b$ = slope of regression line, $= \sum XY / \sum X^2$
$y_i$ = value of $y$ from data at $x = x_i$
$\hat{y}_i$ = value of $y$ from regression line at $x = \hat{x}_i$; note that $\hat{y}_i = a + b\hat{x}_i$ for a straight line

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

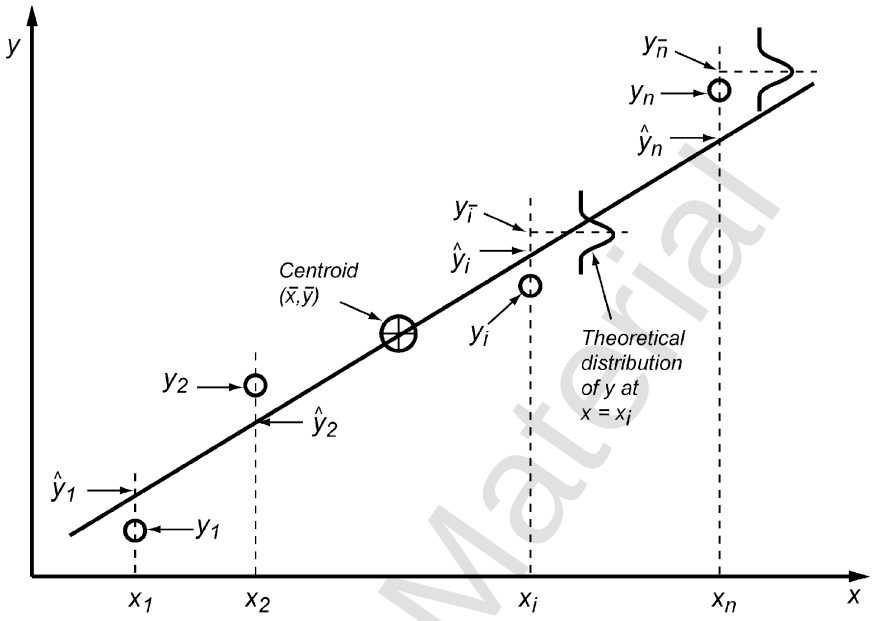



**Figure 3** Schematic of a regression line.

$y_{\bar{i}}$ = mean estimated value of $y_i$ at $x = x_i$; mean of distribution of $y$ values at $x = x_i$; if there is only one measurement of $y$ at $x = x_i$, then that value of $y$ (i.e., $y_i$) is best estimate of $y_{\bar{i}}$

$\nu$ = degrees of freedom in fitting regression line to data ($\nu = n - 2$ for straight line)

$\sum(y_i - \hat{y}_i)^2/\nu = \hat{\sigma}^2_{y,x}$ = unexplained variance (for regression line) where $\hat{\sigma}_{y,x}$ is standard deviation of estimate

$\sigma^2_{\bar{y},x} = \sigma^2_{y,x}/n$ from central limit theorem

$\sigma^2_b = \sigma^2_{y,x}/\sum X^2$ = estimate of variance on slope

***Slope-Centroid Approximation***

This method assumes that the placement uncertainty of the regression line is due to uncertainties in the centroid $(\bar{x}, \bar{y})$ of the data and the slope $b$ of the regression line passing through this centroid. These uncertainties are determined from the following relations and are shown in Fig. 4:

$$\text{Centroid:} \qquad \mu_{\bar{y}} = \hat{\mu}_{\bar{y}} \pm t(\nu,\gamma)\hat{\sigma}_{\bar{y},x} = \bar{y} \pm t\frac{\hat{\sigma}_{y,x}}{\sqrt{n}} \tag{45}$$

$$\text{Slope:} \qquad \mu_b = \hat{b} \pm t(\nu,\gamma)\hat{\sigma}_b = \hat{b} \pm t\frac{\hat{\sigma}_{y,x}}{\sum X^2} \tag{46}$$

***Point-by-Point Approximation***

This is a better approximation than the slope-centroid technique. It gives confidence limits of the points $y_{\bar{i}}$, where

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

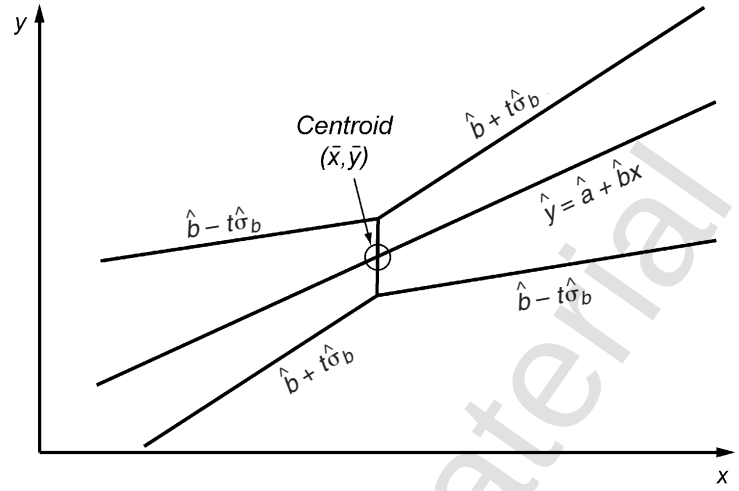



**Figure 4** Schematic illustrating confidence limits on a regression line.

$$\mu_{y\bar{i}} = \hat{y}_{\bar{i}} \pm t(\nu,\gamma)\hat{\sigma}_{y\bar{i}} \tag{47}$$

and

$$\hat{\sigma}^2_{y_i} = \hat{\sigma}^2_{y,x}\left(\frac{1}{n} + \frac{X_i^2}{\sum X_i^2}\right) \tag{48}$$

At the centroid where $X_i = 0$, Eq. (47) reduces to the result for $\mu_{\bar{y}}$ given in Eq. (45).

***Line as a Whole***

More uncertainty is involved when all points for the line are taken collectively. The "price" paid for the additional uncertainty is given by replacing $t(v,\gamma)$ in the confidence interval relation by $\sqrt{2F}$, where $F = F\ (2,\ \nu,\ \gamma)$ obtained from an "$F$" table statistical distribution. Thus the uncertainty interval for placement of the "line as a whole" confidence limit at $x = x_i$ is found by the formula

$$\mu_{\text{Line}} = \hat{y}_i \pm \sqrt{2F}\hat{\sigma}_{y,x}\sqrt{\frac{1}{n} + \frac{X_i}{\sum X_i^2}} \tag{49}$$

***Future Estimate of a Single Point at $x = x_i$***

This gives the expected confidence limits on a future estimate value of $y$ at $x = x_i$ in a relation to a prior regression estimate. This confidence limit can be found from the relation

$$\mu_y = \hat{y}_i \pm t(\nu,\gamma)\hat{\sigma}_{yi} \tag{50}$$

where $\hat{y}_i$ is the best estimate of $y$ from the regression line and

$$\hat{\sigma}^2_{y_i} = \hat{\sigma}^2_{y,x}\left(1 + \frac{1}{n} + \frac{X_i^2}{\sum X_i^2}\right) \tag{51}$$

If one uses $\gamma = 0.99$, nearly all (99%) of the observed data points should be within the uncertainty limits calculated by Eq. (51).

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

**Example 6 Confidence Limits on a Regression Line.** Calibration data for a copper–constantan thermocouple are given:

Temperature: $x = [0, 10, 20, 30, 40\ 50, 60, 70, 80, 90, 100]$ °C

Voltage: $y = [-0.89, -0.53, -9.15, -0.20, 0.61, 1.03, 1.45, 1.88, 2.31, 2.78, 3.22]$ mV

If the variation of $y$ is expected to be linear over the range of $x$ and the uncertainty in temperature is much less than the uncertainty in voltage, then:

1. Determine the linear relation between $y$ and $x$.
2. Test the goodness of fit at the 95% confidence level.
3. Determine the 90% confidence limits of points on the line by the slope-centroid technique.
4. Determine the 90% confidence limits on the intercept of the regression line.
5. Determine the 90% confidence limits on a future estimated point of temperature at 120°C.
6. Determine the 90% confidence limits on the whole line.
7. How much data would be required to determine the centroid of the voltage values within 1% at the 90% confidence level?

The calculations are as follows:

$$\bar{x} = 50.0 \qquad \bar{y} = 1.0827 \qquad y = a + bx$$

$$\Sigma x_i = 550 \qquad \Sigma y_i = 11.91 \qquad \Sigma x_i y_i = 1049.21$$

$$\Sigma (x_i)^2 = 38{,}500.0 \qquad \Sigma (y_i)^2 = 31.6433$$

$$\Sigma (x_i - \bar{x})^2 = \Sigma X_i^2 = 11{,}000.00 \qquad \Sigma (y_i - \bar{y})^2 = \Sigma Y_i^2 = 18.7420$$

$$\Sigma (x_i - \bar{x})(y_i - \bar{y}) = \Sigma X_i Y_i = 453.70$$

$$\Sigma (y_i - \hat{y}_i)^2 = 0.0299$$

**1.** $b = \Sigma XY / \Sigma X^2 = 453.7/11{,}000 = 0.0412$

$a = \bar{y} - b\bar{x} = 1.0827 - (0.0412)(50.00) = 1.0827 - 2.0600$

$= -0.9773$

**2.** $r_{\text{exp}} = \sqrt{\Sigma (\hat{y} - \bar{y})^2 / \Sigma (y_i - \bar{y})^2} = \sqrt{1 - \Sigma (y_i - \hat{y}_i)^2 / \Sigma (y_i - \bar{y})^2}$

$= \Sigma XY / \sqrt{\Sigma X^2 \Sigma Y^2} = 0.998$

$r_{\text{Table}} = r(\alpha, \nu) = r(0.05{,}9) = 0.602$

$P[r_{\text{exp}} > r_{\text{Table}}] = \alpha$ with $H_0$ of no correlation; therefore reject $H_0$ and infer significant correlation

**3.** $t = t(\gamma, \upsilon) = t(0.90{,}9) = 1.833$ (see Ref. 6)

$\hat{\sigma}^2_{y.x} = \Sigma (y_i - \hat{y}_i)^2 / \nu = 0.0299/9 = 0.00333$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$\hat{\sigma}_{y,x} = \pm 0.05777,\ \hat{\sigma}_{\bar{y},x} = \pm 0.0173$$

$$\hat{\sigma}_b = \hat{\sigma}_{y,x}/\sqrt{\Sigma X^2} = \pm 0.000550$$

$$\mu_{\bar{y}} = \hat{\mu}_{\bar{y}} \pm t\hat{\sigma}_{\bar{y},x} = y \pm t\hat{\sigma}_{\bar{y},x}$$

$$= 1.0827 \pm (1.833)(0.01737) = 1.0827 \pm 0.0318$$

$$\mu_b = \hat{b} \pm t\hat{\sigma}_b = 0.0412 \pm (1.833)(0.00055)$$

$$= 0.0412 \pm 0.00101$$

**4.** $$\mu_{y\bar{i}} = \hat{\mu}_{y\bar{i}} \pm t\hat{\sigma}_{y\bar{i}} = y_{\bar{i}} \pm t\hat{\sigma}_{y,x}\sqrt{1/n + X^2/\Sigma X^2}$$

$$= -0.9773 \pm (1.833)(0.0325)$$

$$= -0.9773 \pm 0.0596$$

**5.** $$\mu_{yi} = \hat{\mu}_{yi} \pm t\hat{\sigma}_{yi} = \hat{y}_i + t\hat{\sigma}_{y,x}\sqrt{1 + 1/n + X^2/\Sigma X^2}$$

$$= (-0.9773 + 4.9500) \pm (1.833)(0.0715)$$

$$= 3.9727 \pm 0.1310$$

**6.** $\mu_{y\bar{i}\ \text{Line}} = \hat{\mu}_{y\bar{i}} \pm \sqrt{2F(2, v)}\hat{\sigma}_{y\bar{i}}$; for any given point compare with 4 above

$$= \hat{y}_{\bar{i}} \pm \sqrt{(2)(3.01)}\hat{\sigma}_{y\bar{i}}$$

$$= \hat{y}_{\bar{i}} \pm (2.46)\hat{\sigma}_{y\bar{i}}$$

$$= -0.9773 \pm 0.0799;\ \text{for point of 4 above}$$

**7.** $$\mu_{\bar{y}i} = \hat{\mu}_y \pm t\hat{\sigma}_{\bar{y},x} = \bar{y} \pm t\frac{\hat{\sigma}_{y,x}}{\sqrt{n}}$$

$$\sigma_{\bar{y},x} = \sigma_{\bar{y},x} = \sigma_{y,x}/\sqrt{n}$$

$$\hat{\sigma}_{y,x} = \pm 0.0577$$

$$\mu_{\bar{y}} - \bar{y} = (1\%)(\mu_y) \cong (1\%)(\bar{y}) = 0.010827$$

$$t/\sqrt{n} = (1\%)\bar{y}/\hat{\sigma}_{y,x} = 0.010827/0.05770 = 0.188$$

From the $t$ table[9] at $t/\sqrt{n} = 0.188$ with $\nu = n - 2$ for a straight line (two constraints).

| $\nu$ | $t(0.10,\nu)$ | $t/\sqrt{n}$ |
|---|---|---|
| 60 | 1.671 | 0.213 |
| 90 | 1.663 | 0.174 |
| 75 | 1.674 | 0.188 |

Therefore $n = \nu + 2 = 77$ is the amount of data to obtain to assure the precision and confidence level desired.

## 3.11 Inference and Comparison

Events are not only deterministic but probabilistic. Under certain specified conditions some events will always happen. Some other events, however, may or may not happen. Under the same specified conditions, the latter ones depend on chance and, therefore, the probability of occurrence of such events is of concern. For example, it is quite certain that a tossed unbiased die will fall down. However, it is not at all certain which face will appear on top

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

when the die comes to rest. The probabilistic nature of some events is apparent when questions of the following type are asked. Does medicine *A* cure a disease better than medicine *B*? What is the ultimate strength of 1020 steel? What total mileage will brand *X* tire yield? Which heat treatment process is better for a given part?

Answering such questions involves designing experiments, performing measurements, analyzing the data, and interpreting the results. In this endeavor two common phenomena are observed: (1) repeated measurements of the same attribute differ due to measurement error and resolving capability of the measurement system and (2) corresponding attributes of identical entities differ due to material differences, manufacturing tolerances, tool wear, and so on.

Conclusions based on experiments are statistical inferences and can only be made with some element of doubt. Experiments are performed to make statistical inferences with minimum doubt. Therefore, experiments are designed specifying the data required, amount of data needed, and the confidence limits desired in drawing conclusions. In this process an instrumentation system is specified, a data-taking procedure is outlined, and a statistical method is used to make conclusions at preselected confidence levels.

In statistical analysis of experimental data, the descriptive and inference tasks are considered. The descriptive task is to present a comprehensible set of observations. The inference task determines the truth of the whole by examination of a sample. The inference task requires sampling, comparison, and a variety of statistical tests to obtain unbiased estimates and confidence limits to make decisions.

***Statistical Testing***

A statistical hypothesis is an assertion relative to the distribution of a random variable. The test of a statistical hypothesis is a procedure to accept or reject the hypothesis. A hypothesis is stated such that the experiment attempts to nullify the hypothesis. Therefore, the hypothesis under test is called the *null* hypothesis and symbolized by $H_0$. All alternatives to the null hypothesis are termed the *alternate* hypothesis and are symbolized by $H_1$.[15]

If the results of the experiment cannot reject $H_0$, the experiment cannot detect the differences in measurements at the chosen probability level.

Statistical testing determines if a process or item is better than another with some stated degree of confidence. The concept can be used with a certain statistical distribution to determine the confidence limits.

The following procedure is used in statistical testing:

1. Define $H_0$ and $H_1$.
2. Choose the confidence level $\gamma$ of the test.
3. Form an appropriate probabilistic statement.
4. Using the appropriate statistical distribution, perform the indicated calculation.
5. Make a decision concerning the hypothesis and/or determine confidence limits.

Two types of error are possible in statistical testing. A Type I error is that of rejecting a true null hypothesis (rejecting truth). A Type II error is that of accepting a false null hypothesis (embracing fiction). The confidence levels and sample size are chosen to minimize the probability of making a Type I or Type II error. Figure 5 illustrates the Type I ($\alpha$) and Type II ($\beta$) errors, where

$H_0$ = sample with $n = 1$ comes from $N(10,16)$
$H_1$ = sample with $n = 1$ comes from $N(17,16)$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

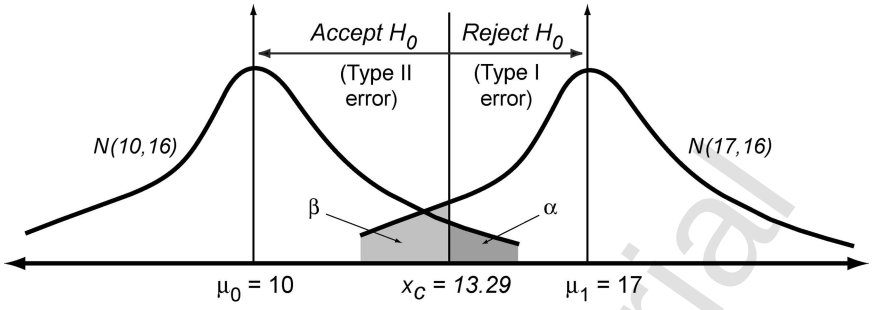



**Figure 5** Probability of Type I and Type II errors.

$\alpha$ = probability of Type I error [concluding data came from $N(17,16)$ when data actually came from $N(10,16)$]
$\beta$ = probability of a Type II error [accepting that data come from $N(10,16)$ when data actually come from $N(17,16)$]

When $\alpha = 0.05$ and $n = 1$, the critical value of $x$ is 16.58 and $\beta = 0.458$. If the value of $x$ obtained from the measurement is less than 16.58, accept $H_0$. If $x$ is larger than 16.58, reject $H_0$ and infer $H_1$. For $\beta = 0.458$ there is a large chance for a Type II error. To minimize this error, increase $\alpha$ or the sample size. For example, when $\alpha = 0.05$ and $n = 4$, the critical value of $x$ is 13.29 and $\beta = 0.032$ as shown in Fig. 5. Thus, the chance of a Type II error is significantly decreased by increasing the sample size. Various statistical tests are summarized in the flow chart shown in Fig. 6.

***Comparison of Variability***

To test whether two samples are from the same population, their variability or dispersion characteristics are first compared using $F$-statistics.[9,15]

If $x$ and $y$ are random variables from two different samples, the parameters $U = \sum(x_i - \bar{x})^2/\sigma_x^2$ and $V = \sum(y_i - \bar{y})^2/\sigma_y^2$ are also random variables and have chi-square distributions (see Fig. 7*a*) with $\nu_1 = n_1 - 1$ and $\nu_2 = n_2 - 1$ degrees of freedom, respectively. The random variable $W$ formed by the ratio $(U/\nu_1)/(V/\nu_2)$ has an $F$-distribution with $\nu_1$ and $\nu_2$ degrees of freedom [i.e., $W \sim F(\gamma, \nu_1, \nu_2)$]. The quotient $W$ is symmetric; therefore, $1/W$ also has an $F$-distribution with $\nu_2$ and $\nu_1$ degrees of freedom [i.e., $1/W \sim F(\alpha, \nu_2, \nu_1)$]. Figure 7*b* shows the $F$-distribution and its probabilistic statement:

$$P[F_L < W < F_R] = \gamma \tag{52}$$

where $W = (U/\nu_1)/(V/\nu_2) = \hat{\sigma}_1^2/\hat{\sigma}_2^2$ with

$$H_0 \text{ as } \sigma_1^2 = \sigma_2^2 \qquad \text{and} \qquad H_1 \text{ as } \sigma_1^2 \neq \sigma_2^2 \tag{53}$$

Here, $W$ is calculated from values of $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ obtained from the samples.

**Example 7 Testing for Homogeneous Variances.** Test the hypothesis at the 90% level of confidence that $\sigma_1^2 = \sigma_2^2$ when the samples of $n_1 = 16$ and $n_2 = 12$ yielded $\hat{\sigma}_1^2 = 5.0$ and $\hat{\sigma}_2^2 = 2.5$. Here $H_0$ is $\sigma_1 = \sigma_2$ and $H_1$ is $\sigma_1 \neq \sigma_2$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

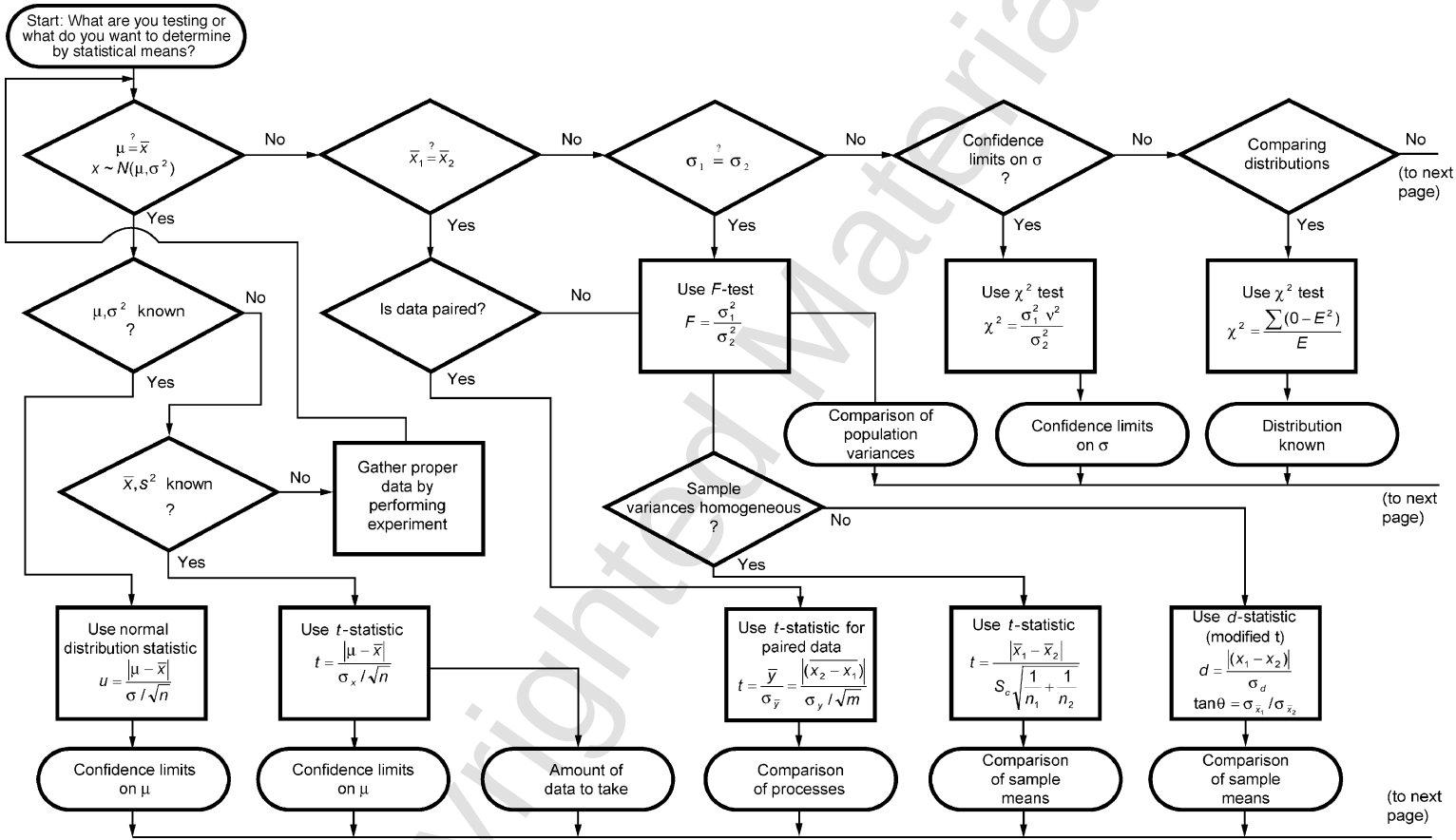



**Figure 6** User's flow chart for statistical tests of data.



Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

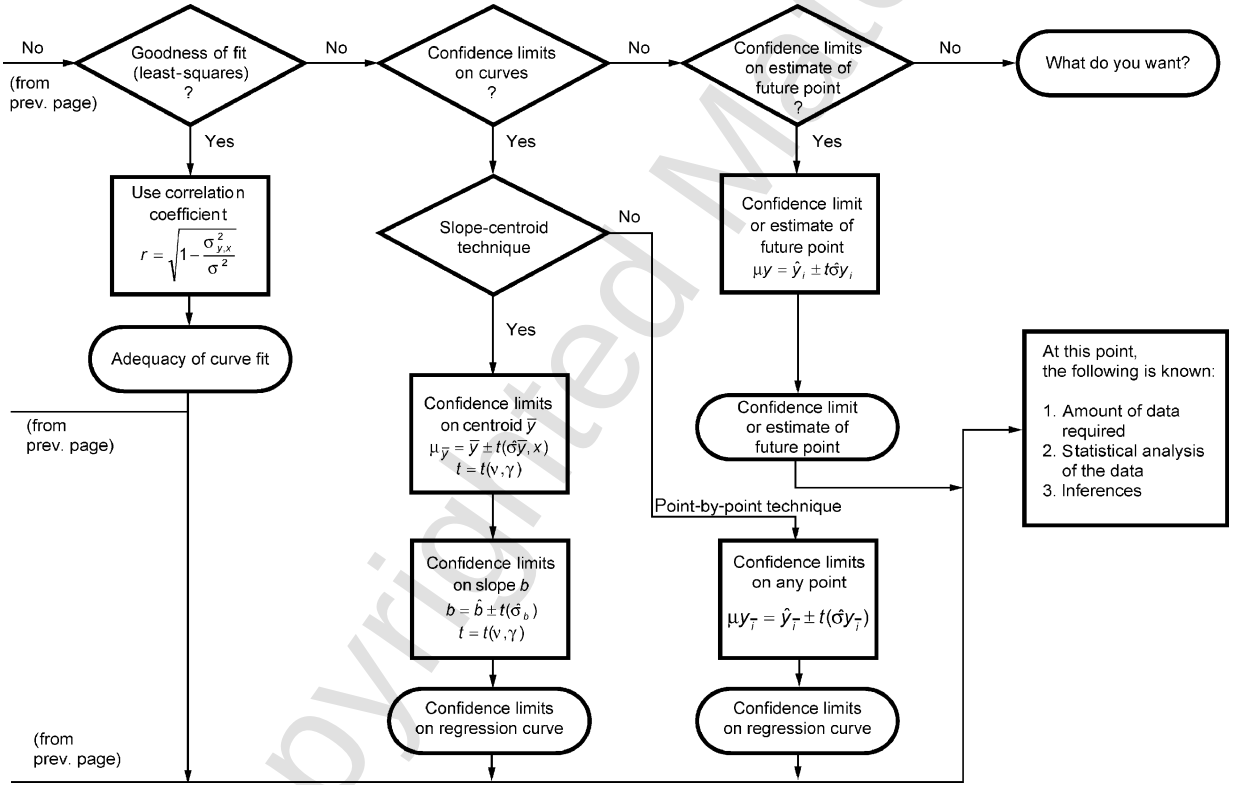



**Figure 6** (*Continued*)

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

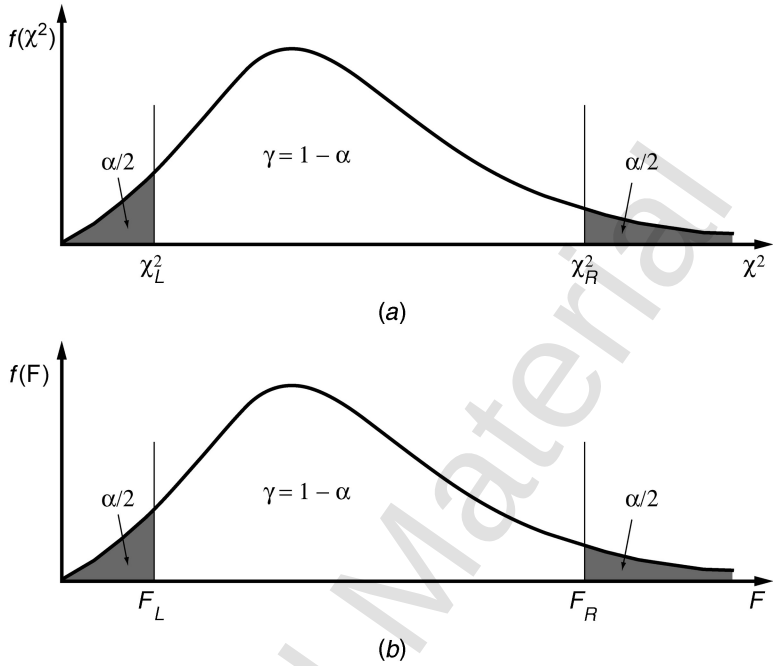



**Figure 7** (*a*) Chi-square distribution and (*b*) *F*-distribution. Also shown are probabilistic statements of the distributions such as that shown in Eq. (52).

$$P[F_L(\nu_1,\nu_2) < \left(\frac{\hat{\sigma}_1}{\hat{\sigma}_2}\right)^2 < F_R(\nu_1,\nu_2)] = \gamma$$

$$P[F_{0.95}(15,11) < \left(\frac{\hat{\sigma}_1}{\hat{\sigma}_2}\right)2 < F_{0.05}(15,11)] = 0.90$$

$$F_{0.95}(15,11) = \frac{1}{F_{0.05}(11,\,15)} = \frac{1}{2.51} = 0.398 \qquad P[0.398 < \left(\frac{\hat{\sigma}_1}{\hat{\sigma}_2}\right)^2 < 2.72] = 0.90$$

Since there was a probability of 90% of $\hat{\sigma}_1^2/\hat{\sigma}_2^2$ ranging between 0.398 and 2.72 and with the actual value of $\hat{\sigma}_1^2/\hat{\sigma}_2^2 = 2.0$, $H_0$ cannot be rejected at the 90% confidence level.

***Comparison of Means***

Industrial experimentation often compares two treatments of a set of parts to determine if a part characteristic such as strength, hardness, or lifetime has been improved. If it is assumed that the treatment does not change the variability of items tested ($H_0$), then the *t*-distribution determines if the treatment had a significant effect on the part characteristics ($H_1$). The *t*-statistic is

$$t = \frac{d - \mu_d}{\sigma_d} \tag{54}$$

where $d = \bar{x}_1 - \bar{x}_2$ and $\mu_d = \mu_1 - \mu_2$. From the propagation of variance, $\sigma_d^2$ becomes

$$\sigma_d^2 = \sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \tag{55}$$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

where $\sigma_1^2$ and $\sigma_2^2$ are each estimates of the population variance $\sigma^2$. A better estimate of $\sigma^2$ is the combined variance $\sigma_c^2$, and it replaces both $\sigma_1^2$ and $\sigma_2^2$ in Eq. (55). The combined variance is determined by weighting the individual estimates of variance based on their degrees of freedom according to the relation

$$\sigma_c^2 = \frac{\hat{\sigma}_1^2 v_1 + \hat{\sigma}_2^2 v_2}{v_1 + v_2} \tag{56}$$

Then

$$\sigma_d^2 = \frac{\hat{\sigma}_c^2}{n_1} + \frac{\hat{\sigma}_c^2}{n_2} = \sigma_c^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \tag{57}$$

Under the hypothesis $H_0$ that $\mu_1 = \mu_2$ (no effect due to treatment), the resulting probabilistic statement is

$$P\left[ -t\sigma_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \bar{x}_1 - \bar{x}_2 < t\sigma_c \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right] = \gamma \tag{58}$$

If the variances of the items being compared are not equal (homogeneous), a modified $t$- (or $d$-) statistic is used,[9,15] where $d$ depends on confidence level $\gamma$, degrees of freedom $v$, and a parameter $\theta$ that depends on the ratio of standard deviations according to

$$\tan \theta = \frac{\hat{\sigma}_1 / \sqrt{n_1}}{\hat{\sigma}_2 / \sqrt{n_2}} \tag{59}$$

The procedure for using the $d$-statistic is the same as described for the $t$-statistic.

**Example 8 Testing for Homogeneous Means.** A part manufacturer has the following data:

| Sample Number | Number of Parts | Mean Lifetime (h) | Variance (h) |
|---|---|---|---|
| 1 | 15 | 2530 | 490 |
| 2 | 11 | 2850 | 360 |

Determine if the lifetime of the parts is due to chance at a 10% significance level (90% confidence level).

Check variance first ($H_0$—homogeneous variances and $H_1$—nonhomogeneous variances):

$$\hat{\sigma}_1^2 = 490\left(\frac{15}{14}\right) = 525$$

$$\hat{\sigma}_2^2 = 360\left(\frac{11}{10}\right) = 396$$

$$\sigma_c^2 = \frac{(14)(525) + (10)(396)}{24} = 471$$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$P[F_L < F_{\text{exp}} < F_R] = \gamma \qquad F_{\text{exp}} = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} = \frac{525}{396} = 1.33$$

$$F_R = F_{0.05}(\nu_1, \nu_2) = F_{0.05}(14,10) = 2.8$$

$$F_L = F_{0.95}(\nu_1, \nu_2) = \frac{1}{F_{0.05}(10,14)} = \frac{1}{2.60} = 0.385$$

Therefore, accept $H_0$ (variances are homogeneous).

Check means next ($H_0$ is $\mu_1 = \mu_2$ and $H_1$ is $\mu_2 \neq \mu_1$):

$$P[-t < t_{\text{calc}} < t] = \gamma$$

$$t = t(\nu_1 + \nu_2, \gamma) = t(24,0.90) = 1.711$$

Therefore, $P[-1.71 < t_{\text{exp}} < 1.71] = 0.90$:

$$t_{\text{exp}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\sigma_c^2\,(1/n_1 + 1/n_2)}}$$

$$= \frac{|2530 - 2850|}{\sqrt{(471)(0.1576)}}$$

$$= \frac{|-320|}{\sqrt{73.6}} = \frac{320}{8.58}$$

$$= 32.3$$

Therefore, reject $H_0$ and accept $H_1$ of $\mu_1 \neq \mu_2$ and infer differences in samples are due not to chance alone but to some real cause.

***Comparing Distributions***

A chi-square distribution is also used for testing whether or not an observed phenomenon fits an expected or theoretical behavior.[15] The chi-square statistic is defined as

$$\sum \left[ \frac{(O_j - E_j)^2}{E_j} \right]$$

where $O_j$ is the observed frequency of occurrence in the $j$th class interval and $E_j$ is the expected frequency of occurrence in the $j$th class interval. The expected class frequency is based on an assumed distribution or a hypothesis $H_0$. This statistical test is used to compare the performance of machines or other items. For example, lifetimes of certain manufactured parts, locations of hole center lines on rectangular plates, and locations of misses from targets in artillery and bombing missions follow chi-square distributions.

The probabilistic statements depend on whether a one-sided test or a two-sided test is being performed.[16] The typical probabilistic statements are

$$P[\chi_{\text{exp}}^2 > \chi^2(\alpha, \nu)] = \alpha \tag{60}$$

for the one-sided test and

$$P[\chi_L^2 < \chi_{\text{exp}}^2 < \chi_R^2] = \gamma \tag{61}$$

for the two-sided test, where

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

$$\chi^2_{\text{exp}} = \frac{\Sigma(O_j - E_j)^2}{E_j} \tag{62}$$

When using Eq. (62) at least five items per class interval are normally used, and a continuity correction must be made if less than three class intervals are used.

**Example 9 Use of the Chi-Square Distribution.** In a test laboratory the following record shows the number of failures of a certain type of manufactured part.

| | Number of Parts | |
|---|---|---|
| Number of Failures | Observed | Expected[a] |
| 0 | 364 | 325 |
| 1 | 376 | 396 |
| 2 | 218 | 243 |
| 3 | 89 | 97 |
| 4 | 33 | 30 |
| 5 | 13 | 7 |
| 6 | 4 | 1 |
| 7 | 3 | 0 |
| 8 | 2 | 0 |
| 9 | 1 | 0 |
| | $\Sigma = 1103$ | |

[a]From an assumed statistical model that the failures are random.

At a 95% confidence level, determine whether the failures are attributable to chance alone or to some real cause. For $H_0$, assume failures are random and not related to a cause so that the expected distribution of failures follows the Poisson distribution. For $H_1$, assume failures are not random and are cause related. The Poisson probability is $P_r = e^{-\mu}\mu^r/r!$, where $\mu = \Sigma x_i P_i$:

$$\mu = (0)\left(\frac{364}{1103}\right) + (1)\left(\frac{376}{1103}\right) + (2)\left(\frac{218}{1103}\right) + (3)\left(\frac{89}{1103}\right) + (4)\left(\frac{33}{1103}\right) + (5)\left(\frac{13}{1103}\right) + \frac{24}{1103} + \frac{21}{1103} + \frac{16}{1103} + \frac{9}{1103}$$

$$= \frac{1346}{1103} = 1.22$$

Then $P_0 = 0.295$ and $E_0 = P_0 n = 325$. Similarly $P_1 = 0.360$, $P_2 = 0.220$, $P_3 = 0.088$, . . . and, correspondingly, $E_1 = 396$, $E_2 = 243$, $E_3 = 97$, . . . are tabulated above for the expected number of parts to have the number of failures listed.

Goodness-of-fit test—$H_0$ represents random failures and $H_1$ represents real cause:

$$P[\chi^2_{\text{exp}} > \chi^2_{\text{Table}}] = \alpha = 0.05$$

$$\chi^2_{\text{Table}} = \chi^2(\nu, \alpha) = \chi^2(5,0.05) = 11.070$$

$$\chi^2_{\text{exp}} = \frac{\Sigma(0 - E)^2}{E}$$

Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com

| $O$ | $E$ | $O-E$ | $(O-E)^2$ | $(O-E)^2/E$ |
|---|---|---|---|---|
| 364 | 325 | 39 | 1520 | 4.6800 |
| 376 | 396 | −20 | 400 | 1.0100 |
| 218 | 243 | −25 | 625 | 2.5700 |
| 89 | 97 | −8 | 64 | 0.6598 |
| 33 | 30 | 3 | 9 | 0.3000 |
| 23 | 8 | 15 | 225 | 28.1000 |

$$\chi^2_{\text{exp}} = 37.320$$

Since there is only a 5% chance of obtaining a calculated chi-square statistic as large as 11.07 under the hypothesis chosen, the null hypothesis is rejected and it is inferred that some real (rather than random) cause exists for the failures with a 95% probability or confidence level.

## REFERENCES


1. E. J. Minnar (ed.), *ISA Transducer Compendium,* Instrument Society of America/Plenum, New York, 1963.
2. "Electrical Transducer Nomenclature and Terminology," ANSI Standard MC 6.1-1975 (ISA S37.1), Instrument Society of America, Research Triangle Park, NC, 1975.
3. Anonymous (ed.), *Standards and Practices for Instrumentation,* Instrument Society of America, Research Triangle Park, NC, 1988.
4. E. O. Doebelin, *Measurement Systems: Application and Design,* 4th ed., McGraw-Hill, New York, 1990.
5. J. G. Webster, *The Measurement, Instrumentation, and Sensors Handbook,* CRC Press in cooperation with IEEE, Boca Raton, FL, 1999.
6. R. S. Figliola and D. E. Beasley, *Theory and Design for Mechanical Measurements,* 3rd ed., Wiley, New York, 2000.
7. C. L. Nachtigal (ed.), *Instrumentation and Control: Fundamentals and Applications,* Wiley, New York, 1990.
8. I. B. Gertsbakh, *Measurement Theory for Engineers,* Springer, New York, 2003.
9. J. B. Kennedy and A. M. Neville, *Basic Statistical Methods for Engineers and Scientists,* 3rd ed., Harper & Row, New York, 1986.
10. A. G. Worthing and J. Geffner, *Treatment of Experimental Data,* Wiley, New York, 1943.
11. C. R. Mischke, *Mathematical Model Building* (*an Introduction to Engineering*), 2nd ed., Iowa State University Press, Ames, IA, 1980.
12. C. Lipson and N. J. Sheth, *Statistical Design and Analysis of Engineering Experiments,* McGraw-Hill, New York, 1973.
13. B. Ostle and R. W. Mensing, *Statistics in Research,* 3rd ed., Iowa State University Press, Ames, IA, 1975.
14. D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers,* Wiley, New York, 2002.
15. S. B. Vardeman, *Statistics for Engineering Problem Solving,* PWS Publishing, Boston, 1994.
16. R. E. Walpole, S. L. Myers, R. H. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists,* 7th ed., Prentice-Hall, Englewood Cliffs, NJ, 2002.


Copyrighted Material

Copyright © 2006 John Wiley & Sons Retrieved from: www.knovel.com