CHAPTER 4

# Statistical Analysis of Measurements Subject to Random Errors

Chapter Outline



## 4.1 Introduction

Having introduced the subject of measurement uncertainty in Chapter 2 and its effect on the quality and accuracy of measurements, the last chapter went on to consider the subject

Measurement and Instrumentation. http://dx.doi.org/10.1016/B978-0-12-800884-3.00004-6

of measurement errors in more detail. We learned in the last chapter that errors in measurement systems can be divided into two classes: systematic and random (precision) errors. Furthermore we learned that the methods for quantifying these two types of errors are fundamentally different. In the case of systematic errors, it is necessary to know first the magnitudes of the physical effects that are affecting a measurement system, but then the measurement error can be estimated as an absolute value. However, in the case of random errors, where the error is caused by random and unpredictable effects, it is only possible to quantify the error in probabilistic terms.

The appropriate means of quantifying and reducing systematic errors were covered in the last chapter. This chapter goes on to examine the nature of random errors in more detail and to explain the various means available for quantifying random errors. This discussion starts with an explanation of now the approximate value of a measurement subject to random errors can be calculated as a mean or median value. The discussion then goes on to introduce the concepts of quantifying the spread of measurements about the mean value as either the standard deviation or the variance. Following this, graphical techniques for displaying the spread of random errors are introduced and, particularly, the Gaussian distribution. This includes an explanation of Gaussian ($z$-function) tables and how these can be used to quantify this spread in statistical terms. The chapter then goes on to explain various other statistical means of analyzing measurement subject to random errors including the standard error of the mean, the chi-squared distribution, goodness of fit tests, identification of rogue data points (data outliers), and the Student $t$-distribution.

So far, we have now covered the means of quantifying both systematic and random errors, the former in the last chapter and the latter in the first part of this chapter. However, an important question still remains. If the final output from a measurement system is calculated by combining together two or more measurements of separate physical variables, how should an overall system measurement error be optimally calculated from the error level in the individual system components. This question is answered in the final part of this chapter.

## 4.2 Mean and Median Values

The average value of a set of measurements of a constant quantity can be expressed as either the mean value or the median value. Historically, the median value was easier for a computer to compute than the mean value because the median computation involves a series of logic operations whereas the mean computation requires addition and division. Many years ago, a computer performed logic operations much faster than arithmetic operations, and there were computational speed advantages in calculating average values by computing the median rather than the mean. However, computer power rapidly increased to a point where this advantage disappeared many years ago.

As the number of measurements increases, the difference between the mean and median values becomes very small. However, the average calculated in terms of the mean value is always slightly closer to the correct value of the measured quantity than the average calculated as the median value for any finite set of measurements. Given the loss of any computational speed advantage because of the massive power of modern-day computers, this means that there is now little argument for calculating average values in terms of the median.

For any set of $n$ measurements $x_1, x_2 \ldots x_n$ of a constant quantity, the most likely true value is the *mean* given by

$$x_{\text{mean}} = \frac{x_1 + x_2 + \cdots x_n}{n} \tag{4.1}$$

This is valid for all data sets, where the measurement errors are distributed equally about the zero error value, that is, where the positive errors are balanced in quantity and magnitude by the negative errors.

The *median* is an approximation to the mean that can be written down without having to sum the measurements. The median is the middle value when the measurements in the data set are written down in ascending order of magnitude. For a set of $n$ measurements $x_1, x_2 \ldots x_n$ of a constant quantity, written down in ascending order of magnitude, the median value is given by

$$x_{\text{median}} = x_{n+1}/2 \tag{4.2}$$

Thus, for a set of nine measurements $x_1, x_2 \ldots x_9$ arranged in the order of magnitude, the median value is $x_5$. For an even number of measurements, the median value is midway between the two center values, that is, for 10 measurements $x_1 \ldots x_{10}$, the median value is given by $(x_5 + x_6)/2$.

Suppose that the length of a steel bar is measured by a number of different observers and the following set of 11 measurements are recorded (in millimeter units). We will call this measurement set A.

398 420 394 416 404 408 400 420 396 413 430 (Measurement set A)

Using (4.1) and (4.2), mean = 409.0 and median = 408. Suppose now that the measurements are taken again using a better measuring rule, and with the observers taking more care, to produce the following measurement set B:

409 406 402 407 405 404 407 404 407 407 408 (Measurement set B)

For these measurements, mean = 406.0 and median = 407. Which of the two measurement sets A and B, and the corresponding mean and median values, should we have most confidence in? Intuitively, we can regard measurement set B as being more reliable since the measurements are much closer together. In set A, the spread between the smallest (396) and largest (430) value is 34, while in set B, the spread is only 6.

- *Thus, the smaller the spread of the measurements, the more confidence we have in the mean or median value calculated.*

Let us now see what happens if we increase the number of measurements by extending measurement set B to 23 measurements. We will call this measurement set C.

409 406 402 407 405 404 407 404 407 407 408 406 410
406 405 408 406 409 406 405 409 406 407 (Measurement set C)

Now, mean = 406.5 and median = 406.

- *This confirms our earlier statement that the median value tends toward the mean value as the number of measurements increases.*

## 4.3 Standard Deviation and Variance

Expressing the spread of measurements simply as the range between the largest and smallest value is not in fact a very good way of examining how the measurement values are distributed about the mean value. A much better way of expressing the distribution is to calculate the variance or standard deviation of the measurements. The starting point for calculating these parameters is to calculate the deviation (error) $d_i$ of each measurement $x_i$ from the mean value $x_{\text{mean}}$ in a set of measurements $x_1, x_2 \ldots x_n$:

$$d_i = x_i - x_{\text{mean}} \tag{4.3}$$

The *variance* ($V_s$) of the set of measurements is formally defined as the mean of the squares of the deviations:

$$V_s = \frac{d_1^2 + d_2^2 \cdots d_n^2}{n} \tag{4.4}$$

The *standard deviation* ($\sigma_s$) of the set of measurements is defined as the square root of the variance:

$$\sigma = \sqrt{V_s} = \sqrt{\frac{d_1^2 + d_2^2 \cdots d_n^2}{n}} \tag{4.5}$$

Unfortunately, these formal definitions for the variance and standard deviation of data are made with respect to an infinite population of data values whereas in all practical situations, we can only have a finite set of measurements. We have previously made the observation that the mean value $x_m$ of a finite set of measurements will differ from the true mean $\mu$ of the theoretical infinite population of measurements that the finite set is part of.

This means that there is an error in the mean value $x_{mean}$ used in the calculation of $d_i$ in Eqn (4.3). Because of this, Eqns (4.4) and (4.5) give a biased estimate that tends to underestimate the variance and standard deviation of the infinite set of measurements. A better prediction of the variance of the infinite population can be obtained by applying the Bessel correction factor $(n/n-1)$ to the formula for $V_s$ in Eqn (4.4):

$$V = \left(\frac{n}{n-1}\right)V_s = \frac{d_1^2 + d_2^2 \cdots d_n^2}{n-1} \tag{4.6}$$

where $V_s$ is the variance of the finite set of measurements and $V$ is the variance of the infinite population of measurements.

This leads to a similar better prediction of the standard deviation by taking the square root of the variance in Eqn (4.6):

$$\sigma = \sqrt{V} = \sqrt{\frac{d_1^2 + d_2^2 \cdots d_n^2}{n-1}} \tag{4.7}$$

## Example 4.1

Calculate $\sigma$ and $V$ for measurement sets A, B, and C above.

## Solution

First, draw a table of measurements and deviations for set A (mean = 409 as calculated earlier) as follows:

| Measurement | 398 | 420 | 394 | 416 | 404 | 408 | 400 | 420 | 396 | 413 | 430 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Deviation from mean | −11 | +11 | −15 | +7 | −5 | −1 | −9 | +11 | −13 | +4 | +21 |
| (deviations)$^2$ | 121 | 121 | 225 | 49 | 25 | 1 | 81 | 121 | 169 | 16 | 441 |

$\sum(\text{deviations})^2 = 1370$; $n$ = number of measurements = 11.

Then, from (4.6) and (4.7),
$V = \sum(\text{deviations})^2/n-1; \quad = 1370/10 = 137; \quad \sigma = \sqrt{V} = 11.7.$

The measurements and deviations for set B are (mean = 406 as calculated earlier) as follows:

| Measurement | 409 | 406 | 402 | 407 | 405 | 404 | 407 | 404 | 407 | 407 | 408 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Deviation from mean | +3 | 0 | −4 | +1 | −1 | −2 | +1 | −2 | +1 | +1 | +2 |
| (deviations)$^2$ | 9 | 0 | 16 | 1 | 1 | 4 | 1 | 4 | 1 | 1 | 4 |

From this data, using (4.6) and (4.7), $V = 4.2$ and $\sigma = 2.05$.

The measurements and deviations for set C are (mean = 406.5 as calculated earlier) as follows:

| Measurement | 409 | 406 | 402 | 407 | 405 | 404 | 407 | 404 | 407 | 407 | 408 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Deviation from mean | +2.5 | −0.5 | −4.5 | +0.5 | −1.5 | −2.5 | +0.5 | −2.5 | +0.5 | +0.5 | +1.5 |
| $(\text{deviations})^2$ | 6.25 | 0.25 | 20.25 | 0.25 | 2.25 | 6.25 | 0.25 | 6.25 | 0.25 | 0.25 | 2.25 |

| Measurement | 406 | 410 | 406 | 405 | 408 | 406 | 409 | 406 | 405 | 409 | 406 | 407 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Deviation from mean | −0.5 | +3.5 | −0.5 | −1.5 | +1.5 | −0.5 | +2.5 | −0.5 | −1.5 | +2.5 | −0.5 | +0.5 |
| $(\text{deviations})^2$ | 0.25 | 12.25 | 0.25 | 2.25 | 2.25 | 0.25 | 6.25 | 0.25 | 2.25 | 6.25 | 0.25 | 0.25 |

From these data, using (4.6) and (4.7), $V = 3.53$ and $\sigma = 1.88$.

Note that the smaller values of $V$ and $\sigma$ for measurement set B compared with A correspond with the respective size of the spread in the range between maximum and minimum values for the two sets.

- *Thus, as $V$ and $\sigma$ decrease for a measurement set, we are able to express greater confidence that the calculated mean or median value is close to the true value, that is, that the averaging process has reduced the random error value close to zero.*
- *Comparing $V$ and $\sigma$ for measurement sets B and C, $V$ and $\sigma$ get smaller as the number of measurements increases, confirming that confidence in the mean value increases as the number of measurements increases.*

We have observed so far that random errors can be reduced by taking the average (mean or median) of a number of measurements. However, although the mean or median value is close to the true value, it would only become exactly equal to the true value if we could average an infinite number of measurements. As we can only make a finite number of measurements in a practical situation, the average value will still have some error. This error can be quantified as the *standard error of the mean*, which will be discussed in detail a little later. However, before that, the subject of graphical analysis of random measurement errors needs to be covered.

■

## 4.4 Graphical Data Analysis Techniques—Frequency Distributions

Graphical techniques are a very useful way of analyzing the way in which random measurement errors are distributed. The simplest way of doing this is to draw a *histogram*, in which bands of equal width across the range of measurement values are defined and the number of measurements within each band is counted. The bands are often given the name *data bins*. There are two alternative rules for calculating the best number of data bins to use:

The *Sturges rule* calculates the number of bands as follows:

$$\text{Number of bands} = 1 + 3.3\log_{10}(n),$$

where $n$ is the number of measurement values.

The *Rice rule* calculates the number of bands as $2n^{1/3}$.

Obviously the result produced has to be rounded to the nearest integer in both cases.

When $n$ is relatively small, the two rules suggest the same number of bins. However, for larger values of $n$, the Rice rule calculates a larger number of bins than the Sturges rule. This is summarized in the table below:

| Number of Measurements | Number of Bins Calculated by Sturges Rule | Number of Bins by Sturges (Rounded) | Number of Bins Calculated by Rice Rule | Number of Bins by Rice (Rounded) |
|---|---|---|---|---|
| 10 | 4.3 | 4 | 4.3 | 4 |
| 15 | 4.9 | 5 | 4.9 | 5 |
| 20 | 5.3 | 5 | 5.4 | 5 |
| 25 | 5.6 | 6 | 5.8 | 6 |
| 30 | 5.9 | 6 | 6.2 | 7 |
| 50 | 6.6 | 7 | 7.4 | 7 |
| 100 | 7.6 | 8 | 9.3 | 9 |
| 200 | 8.6 | 9 | 11.7 | 12 |

These rules should be regarded as a good guide but, in any given situation, there may be a good reason for varying the number of bins away from the number suggested. For larger numbers of measurements, it is common practice to use a bin number in between the numbers recommended by the two rules. Special consideration is also needed when dealing with measurement data with random errors. For these, a symmetrical histogram is expected and this is shown better if the number of bins is odd rather than even.

## Example 4.2

Draw a histogram for the 23 measurements in set C of the length measurement data given in Section 4.2.

## Solution

For 23 measurements, the recommended number of bands calculated according to the Sturges rule is $1 + 3.3\log10(23) = 5.49$. This rounds to 5, since the number of bands must be an integer number.

To cover the span of measurements in data set C with 5 bands, the data bands need to be 2-mm wide. The boundaries of these bands must be carefully chosen so that no measurements fall on the boundary between different bands and cause ambiguity about which band to put them in. Since the measurements are integer numbers, this can be easily accomplished by defining the range of the first band as 401.5—403.5 and so on. A histogram can now be drawn as in Figure 4.1 by counting the number of measurements in each band.

In the first band from 401.5 to 403.5, there is just 1 measurement and so the height of the histogram in this band is 1 unit.

In the next band from 403.5 to 405.5 there are 5 measurements and so the height of the histogram in this band is 5 units.

The rest of the histogram is completed in a similar fashion.

■

When a histogram is drawn using a sufficiently large number of measurements, it will have the characteristic shape shown by truly random data, with symmetry about the mean value of the measurements. However, for a relatively small number of measurements, only approximate symmetry in the histogram can be expected about the mean value. It is a matter of judgment as to whether the shape of a histogram is close enough to symmetry to justify a conclusion that the data on which it is based is truly random. It should be noted that the 23 measurements used to draw the histogram in Figure 4.1 were carefully chosen

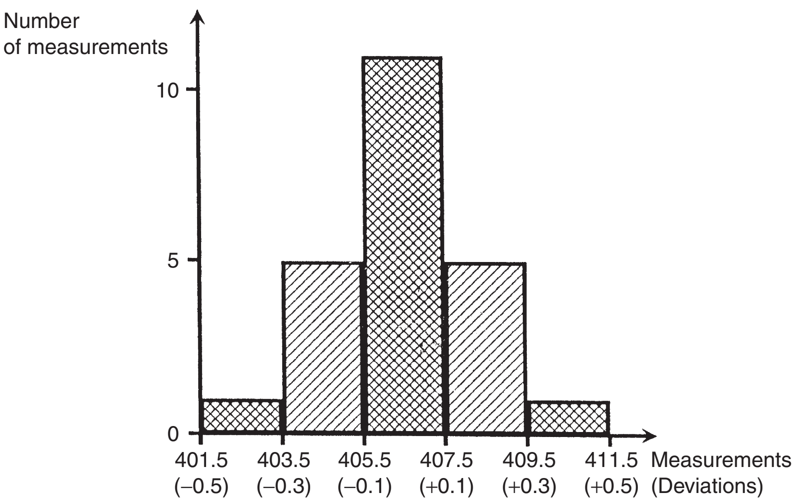


**Figure 4.1**
Histogram of measurements and deviations.

to produce a symmetrical histogram but exact symmetry would not normally be expected for a measurement data set as small as 23.

As it is the actual value of measurement error that is usually of most concern, it is often more useful to draw a histogram of the deviations of the measurements from the mean value rather than to draw a histogram of the measurements themselves. The starting point for this is to calculate the deviation of each measurement away from the calculated mean value. Then a *histogram of deviations* can be drawn by defining deviation bands of equal width and counting the number of deviation values in each band. This histogram has exactly the same shape as the histogram of the raw measurements except that the scaling of the horizontal axis has to be redefined in terms of the deviation values (these units are shown in brackets on Figure 4.1).

Let us now explore what happens to the histogram of deviations as the number of measurements increases. As the number of measurements increases, smaller bands can be defined for the histogram, which retains its basic shape but then consists of a larger number of smaller steps on each side of the peak. In the limit, as the number of measurements approaches infinity, the histogram becomes a smooth curve known as a *frequency distribution curve* as shown in Figure 4.2. The ordinate of this curve is the frequency of occurrence of each deviation value, $F(D)$, and the abscissa is the magnitude of deviation, $D$.

The symmetry of Figures 4.1 and 4.2 about the zero deviation value is very useful for showing graphically that the measurement data only has random errors. Although these

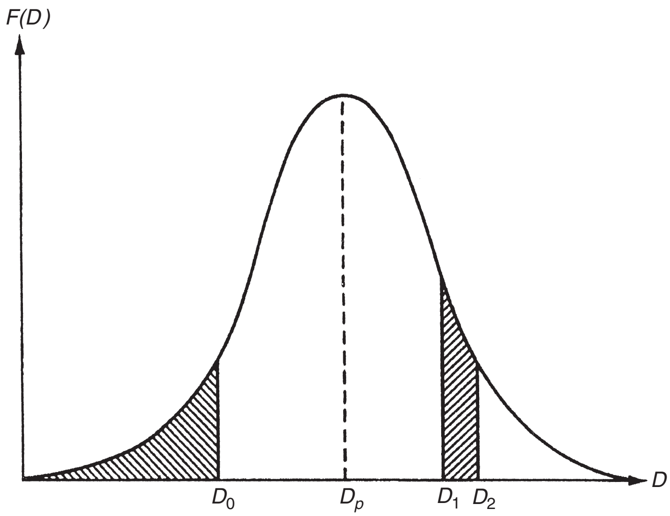


**Figure 4.2**
Frequency distribution curve of deviations.

figures cannot easily be used to quantify the magnitude and distribution of the errors, very similar graphical techniques do achieve this. If the height of the frequency distribution curve is normalized such that the area under it is unity, then the curve in this form is known as a *probability curve*, and the height $F(D)$ at any particular deviation magnitude $D$ is known as the *probability density function* (p.d.f.). The condition that the area under the curve is unity can be expressed mathematically as follows:

$$\int_{-\infty}^{\infty} F(D)\mathrm{d}D = 1$$

The probability that the error in any one particular measurement lies between two levels $D_1$ and $D_2$ can be calculated by measuring the area under the curve contained between two vertical lines drawn through $D_1$ and $D_2$, as shown by the right-hand hatched area in Figure 4.2. This can be expressed mathematically as follows:

$$P(D_1 \le D \le D_2) = \int_{D_1}^{D_2} F(D)\mathrm{d}D \tag{4.8}$$

Of particular importance for assessing the maximum error likely in any one measurement is the *cumulative distribution function* (c.d.f.). This is defined as the probability of observing a value less than or equal to $D_o$, and is expressed mathematically as follows:

$$P(D \le D_0) = \int_{-\infty}^{D_0} F(D)\mathrm{d}D \tag{4.9}$$

Thus, the c.d.f. is the area under the curve to the left of a vertical line drawn through $D_o$, as shown by the left-hand hatched area on Figure 4.2.

The deviation magnitude $D_p$ corresponding with the peak of the frequency distribution curve (Figure 4.2) is the value of deviation that has the greatest probability. If the errors are entirely random in nature, then the value of $D_p$ will be equal to zero. Any nonzero value of $D_p$ indicates systematic errors in the data, in the form of a bias that is often removable by recalibration.

## 4.5 Gaussian (Normal) Distribution

Measurement sets that only contain random errors usually conform to a distribution with a particular shape that is called *Gaussian*, although this conformance must always be tested (see the later section headed "Goodness of fit"). The shape of a Gaussian curve is such that the frequency of small deviations from the mean value is much greater than the frequency of large deviations. This coincides with the usual expectation in measurements subject to random errors that the number of measurements with a small error is much larger than the number of measurements with a large error. Alternative names for the Gaussian distribution are the *Normal distribution* or *Bell-shaped distribution*. A Gaussian curve is formally defined as a normalized frequency distribution that is symmetrical about the line of zero error and in which the frequency and magnitude of quantities are related by the expression

$$F(\boldsymbol{x}) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\left[-(\boldsymbol{x}-m)^2/2\sigma^2\right]} \tag{4.10}$$

where $m$ is the mean value of the data set $x$ and the other quantities are as defined before. Equation (4.10) is particularly useful for analyzing a Gaussian set of measurements and predicting how many measurements lie within some particular defined range. If the measurement deviations $D$ are calculated for all measurements such that $D = x - m$, then the curve of deviation frequency $F(D)$ plotted against deviation magnitude $D$ is a Gaussian curve known as the *error frequency distribution curve*. The mathematical relationship between $F(D)$ and $D$ can then be derived by modifying Eqn (4.10) to give

$$F(D) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\left[-D^2/2\sigma^2\right]} \tag{4.11}$$

The shape of a Gaussian curve is strongly influenced by the value of $\sigma$, with the width of the curve decreasing as $\sigma$ becomes smaller. As a smaller $\sigma$ corresponds with the typical deviations of the measurements from the mean value becoming smaller, this confirms the earlier observation that the mean value of a set of measurements gets closer to the true value as $\sigma$ decreases.

If the standard deviation is used as a unit of error, the Gaussian curve can be used to determine the probability that the deviation in any particular measurement in a Gaussian data set is greater than a certain value. By substituting the expression for $F(D)$ in (4.11) into the probability Eqn (4.8), the probability that the error lies in a band between error levels $D_1$ and $D_2$ can be expressed as

$$P(D_1 \leq D \leq D_2) = \int_{D_1}^{D_2} \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{\left(-D^2/2\sigma^2\right)} \mathrm{d}D \tag{4.12}$$

Solution of this expression is simplified by the substitution

$$z = D/\sigma \tag{4.13}$$

The effect of this is to change the error distribution curve into a new Gaussian distribution that has a standard deviation of one ($\sigma = 1$) and a mean of zero. This new form, shown in

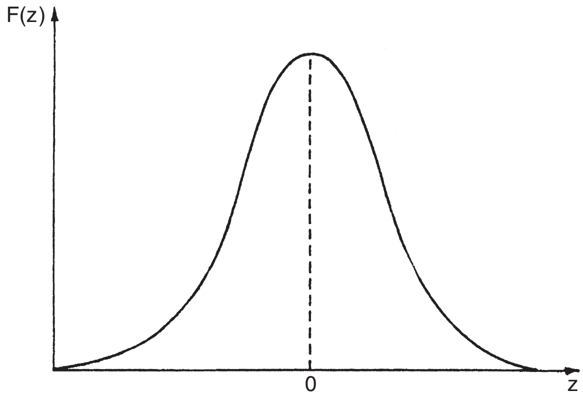


**Figure 4.3**
Standard Gaussian curve ($F(z)$ vs $z$).

Figure 4.3, is known as a *standard Gaussian curve* (or sometimes as a *z-distribution*), and the dependent variable is now $z$ instead of $D$. Equation (4.12) can now be reexpressed as

$$P(D_1 \leq D \leq D_2) = P(z_1 \leq z \leq z_2) = \int_{z_1}^{z_2} \frac{1}{\sigma\sqrt{2\pi}} e^{(-z^2/2)} dz \qquad (4.14)$$

Unfortunately, neither Eqn (4.12) nor (4.14) can be solved analytically using tables of standard integrals, and numerical integration provides the only method of solution. However, in practice, the tedium of numerical integration can be avoided when analyzing data because the standard form of Eqn (4.14), and its independence from the particular values of the mean and standard deviation of the data, means that standard Gaussian tables that tabulate $F(z)$ for various values of $z$ can be used.

## 4.6 Standard Gaussian Tables (z-Distribution)

A standard Gaussian table (sometimes called the *z-distribution*), such as that shown in Table 4.1, tabulates the area under the Gaussian curve $F(z)$ for various values of $z$, where $F(z)$ is given by

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} e^{(-z^2/2)} dz \qquad (4.15)$$

Thus, $F(z)$ gives the proportion of data values that are less than or equal to $z$. This proportion is the area under the curve of $F(z)$ against $z$ that is to the left of $z$. Therefore, the expression given in (4.14) has to be evaluated as $[F(z_2) - F(z_1)]$. Study of Table 4.1 shows that $F(z) = 0.5$ for $z = 0$. This confirms that, as expected, the number of data

**Table 4.1: Error function table (area under a Gaussian curve or *z*-distribution)**

| | $F(z)$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *z* | **0.00** | **0.01** | **0.02** | **0.03** | **0.04** | **0.05** | **0.06** | **0.07** | **0.08** | **0.09** |
| 0.0 | 0.5000 | 0.5000 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7703 | 0.7734 | 0.7764 | 0.7793 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8906 | 0.8925 | 0.8943 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9648 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9924 | 0.9926 | 0.9928 | 0.9930 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9986 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |
| 3.5 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 |
| 3.6 | 0.9998 | 0.9998 | 0.9998 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |

values $\leq 0$ is 50% of the total. This must be so if the data only have random errors. It will also be observed that Table 4.1, in common with most published standard Gaussian tables, only give $F(z)$ for positive values of $z$. For negative values of $z$, we can make use of the following relationship because the frequency distribution curve is normalized:

$$F(-z) = 1 - F(z) \tag{4.16}$$

$F(-z)$ is the area under the curve to the left of $(-z)$, that is, it represents the proportion of data values $\leq -z$.

## Example 4.3

How many measurements in a data set subject to random errors lie outside deviation boundaries of $+\sigma$ and $-\sigma$, that is, how many measurements have a deviation greater than $|\sigma|$?

## Solution

The required number is represented by the sum of the two shaded areas in Figure 4.4. This can be expressed mathematically as $P(E < -\sigma \text{ or } E > +\sigma) = P(E < -\sigma) + P(E > +\sigma)$.

For $E = -\sigma$, $z = -1.0$ (from Eqn (4.11)).

Using Table 4.1 $P(E < -\sigma) = F(-1) = 1 - F(1) = 1 - 0.8413 = 0.1587$.

Similarly, for $E = +\sigma$, $z = +1.0$, Table 4.1 gives

$P(E > +\sigma) = 1 - P(E < +\sigma) = 1 - F(1) = 1 - 0.8413 = 0.1587$. (This last step is valid because the frequency distribution curve is normalized such that the total area under it is unity.)

Thus, $P[E < -\sigma] + P[E > +\sigma] = 0.1587 + 0.1587 = 0.3174 \sim 32\%$, that is, 32% of the measurements lie outside the $\pm\sigma$ boundaries, then 68% of the measurements lie inside.

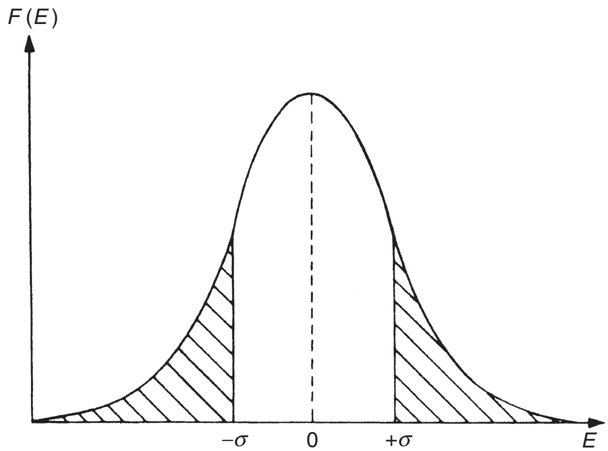


**Figure 4.4**
$\pm\sigma$ boundaries.

The above analysis shows that, for Gaussian-distributed data values, 68% of the measurements have deviations that lie within the bounds of $\pm\sigma$. Similar analysis shows that boundaries of $\pm 2\sigma$ contain 95.4% of data points, and extending the boundaries to $\pm 3\sigma$ encompasses 99.7% of data points. The probability of any data point lying outside particular deviation boundaries can therefore be expressed by the following table.

| Deviation Boundaries | % of Data Points within Boundary | Probability of Any Particular Data Point Being Outside Boundary |
|---|---|---|
| $\pm\sigma$ | 68.0 | 32.0% |
| $\pm 2\sigma$ | 95.4 | 4.6% |
| $\pm 3\sigma$ | 99.7 | 0.3% |

## 4.7 Standard Error of the Mean

The foregoing analysis has examined the way in which measurements with random errors are distributed about the mean value. However, we have already observed that some error exists between the mean value of a finite set of measurements and the true value, that is, averaging a number of measurements will only yield the true value if the number of measurements is infinite. If several subsets are taken from an infinite data population with a Gaussian distribution, then, by the central limit theorem, the means of the subsets will form a Gaussian distribution about the mean of the infinite data set. The standard deviation of the mean values of a series of finite sets of measurements relative to the true mean (the mean of the infinite population that the finite set of measurements is drawn from) is defined as the *standard error of the mean*, $\alpha$. This is calculated as

$$\alpha = \sigma \big/ \sqrt{n} \tag{4.17}$$

Clearly, $\alpha$ tends toward zero as the number of measurements ($n$) in the data set expands toward infinity.

The next question is how do we use the standard error of the mean to predict the error between the calculated mean of a finite set of measurements and the mean of the infinite population? In other words, if we use the mean value of a finite set of measurements to predict the true value of the measured quantity, what is the likely error in this prediction? This likely error can only be expressed in probabilistic terms. All we know for certain is the standard deviation of the error, which is expressed as $\alpha$ in Eqn (4.17). We also know that a range of $\pm$ one standard deviation (i.e., $\pm\alpha$) encompasses 68% of the deviations of sample means either side of the true value. Thus we can say that the measurement value obtained by calculating the mean of a set of n measurements, $x_1, x_2 \ldots x_n$, can be expressed as

$$x = x_{\text{mean}} \pm \alpha$$

with 68% certainty that the magnitude of the error does not exceed $|\alpha|$. For the data set C of length measurements used earlier, $n = 23$, $\sigma = 1.88$, and $\alpha = 0.39$. The length can therefore be expressed as $406.5 \pm 0.4$ (68% confidence limit).

The problem of expressing the error with 68% certainty is that there is a 32% chance that the error is greater than $\alpha$. Such a high probability of the error being greater than $\alpha$ may not be acceptable in many situations. If this is the case, we can use the fact that a range of $\pm$ two standard deviations, that is, $\pm 2\alpha$ encompasses 95.4% of the deviations of sample means either side of the true value. Thus, we can express the measurement value as

$$x = x_{\text{mean}} \pm 2\alpha$$

with 95.4% certainty that the magnitude of the error does not exceed $|2\alpha|$. This means that there is only a 4.6% chance that the error exceeds $2\alpha$. Referring again to set C of length measurements, $2\sigma = 3.76$, $2\alpha = 0.78$ and the length can be expressed as $406.5 \pm 0.8$ (95.4% confidence limits).

If we wish to express the maximum error with even greater probability that the value is correct, we could use $\pm 3\alpha$ limits (99.7% confidence). In this case, for the length measurements again, $3\sigma = 5.64$, $3\alpha = 1.17$ and the length should be expressed as $406.5 \pm 1.2$ (99.7% confidence limits). There is now only a 0.3% chance (3 in 1000) that the error exceeds this value of 1.2.

## Example 4.4

In a practical exercise to determine the freezing point of a metal alloy, the following measurements of the freezing point temperature were obtained:

519.5 521.7 518.9 520.3 521.4 520.1 519.8 520.2 518.6 521.5

Express the mean value and the error boundaries expressed to (a) 68% confidence limits, (b) 95.4% confidence limits, and (c) 99.7% confidence limits.

## Solution

First calculate the mean value of the measurements.

$$\begin{aligned} \text{Mean} &= \sum(519.5\ \ 521.7\ \ 518.9\ \ 520.3\ \ 521.4\ \ 520.1\ \ 519.8\ \ 520.2\ \ 518.6\ \ 521.5) \\ &= 520.2 \end{aligned}$$

Next, calculate the deviations of the measurements from the mean, and hence the standard deviation.

| Measurement | 519.5 | 521.7 | 518.9 | 520.3 | 521.4 | 520.0 | 519.8 | 520.3 | 518.6 | 521.5 |
|---|---|---|---|---|---|---|---|---|---|---|
| Deviation from mean | −0.7 | +1.5 | −1.3 | +0.1 | +1.2 | −0.2 | −0.4 | +0.1 | −1.6 | +1.3 |
| $(\text{deviations})^2$ | 0.49 | 2.25 | 1.69 | 0.01 | 1.44 | 0.04 | 0.16 | 0.01 | 2.56 | 1.69 |

$\sum(\text{deviations})^2 = 10.34$; $n$ = number of measurements = 10.

Then, from (4.6) and (4.7), $\sigma = \sqrt{\frac{\sum(\text{deviations})^2}{n-1}} = 10.34/9 = 1.149$.

The standard error of the mean is given by

$$\alpha = \sigma\Big/\sqrt{n} = 1.149\Big/\sqrt{9} = 0.383$$

The mean of the measurements expressed to 68% confidence limits is given by

$$x = x_{\text{mean}} \pm \alpha = 520.2 \pm 0.4$$

The mean of the measurements expressed to 95.4% confidence limits is given by

$$x = x_{\text{mean}} \pm 2\alpha = 520.2 \pm 0.8$$

The mean of the measurements expressed to 99.7% confidence limits is given by

$$x = x_{\text{mean}} \pm 3\alpha = 520.2 \pm 1.2$$

■

## 4.8 Estimation of Random Error in a Single Measurement

In many situations, where measurements are subject to random errors, it is not practical to take repeated measurements and find the average value. Also, the averaging process becomes invalid if the measured quantity does not remain at a constant value, as is usually the case when process variables are being measured. Thus, if only one measurement can be made, some means of estimating the likely magnitude of error in it is required. The normal approach to this is to calculate the error within 95% confidence limits, that is, to calculate the value of the deviation $D$ such that 95% of the area under the probability curve lies within limits of $\pm D$. These limits correspond to a deviation of $\pm 1.96\sigma$. Thus, it is necessary to maintain the measured quantity at a constant value while a number of measurements are taken in order to create a reference measurement set from which $\sigma$ can be calculated. Subsequently, the maximum likely deviation in a single measurement can be expressed as deviation $= \pm 1.96\sigma$. However, this only expresses the maximum likely deviation of the measurement from the calculated mean of the reference measurement set, which is not the true value as observed earlier. Thus the calculated value for the standard error of the mean has to be added to the likely maximum deviation value. To be consistent, this should be expressed to the same 95% confidence limits. Thus, the maximum likely error in a single measurement can be expressed as

$$\text{Error} = \pm 1.96(\sigma + \alpha) \tag{4.18}$$

## Example 4.5

Suppose that a standard mass is measured 30 times with the same instrument to create a reference data set, and the calculated values of $\sigma$ and $\alpha$ are $\sigma = 0.46$ and $\alpha = 0.08$. If the instrument is then used to measure an unknown mass and the reading is 105.6 kg, how should the mass value be expressed?

## Solution

Using (4.18), $1.96(\sigma + \alpha) = 1.06$. The mass value should therefore be expressed as $105.6 \pm 1.1$ kg.

Before leaving this matter, it must be emphasized that the maximum error specified for a measurement is only specified for the confidence limits defined. Thus, if the maximum error is specified as $\pm 1\%$ with 95% confidence limits, this means that there is still 1 chance in 20 that the error will exceed $\pm 1\%$.

## *4.9 Distribution of Manufacturing Tolerances*

Many aspects of manufacturing processes are subject to random variations caused by factors that are similar to those that cause random errors in measurements. In most cases, these random variations in manufacturing, which are known as *tolerances*, fit a Gaussian distribution, and the previous analysis of random measurement errors can be applied to analyze the distribution of these variations in manufacturing parameters.

## Example 4.6

An integrated circuit chip contains $10^5$ transistors. The transistors have a mean current gain of 20 and a standard deviation of 2. Calculate the following:

(a) the number of transistors with a current gain between 19.8 and 20.2.
(b) the number of transistors with a current gain greater than 17.

## Solution

(a) The proportion of transistors where $19.8 < \text{gain} < 20.2$ is

$$P[X < 20] - P[X < 19.8] = P[z < 0.2] - P[z < -0.2] \quad (\text{for } z = (X - \mu)/\sigma)$$

For $X = 20.2$; $z = 0.1$ and for $X = 19.8$; $z = -0.1$

From tables, $P[z < 0.1] = 0.5398$ and thus $P[z < -0.1] = 1 - P[z < 0.1] = 1 - 0.5398 = 0.4602$.

Hence, $P[z < 0.1] - P[z < -0.1] = 0.5398 - 0.4602 = 0.0796$.

Thus $0.0796 \times 10^5 = 7960$ transistors have a current gain in the range from 19.8 to 20.2.

(b) The number of transistors with gain $> 17$ is given by

$$P[x > 17] = 1 - P[x < 17] = 1 - P[z < -1.5] = P[z < +1.5] = 0.9332$$

Thus, 93.32%, that is, 93,320 transistors have a gain $> 17$.

■

## 4.10 Chi-Squared ($\chi^2$) Distribution

We have already observed the fact that, if we calculate the mean value of successive sets of samples of N measurements, the means of those samples form a Gaussian distribution about the true value of the measured quantity (the true value being the mean of the infinite data set that the set of samples are part of). The standard deviation of the distribution of the mean values was quantified as the standard error of the mean.

It is also useful for many purposes to look at the distribution of the variance of successive sets of samples of $N$ measurements that form part of a Gaussian distribution. This is expressed as the chi-squared distribution $F(\chi^2)$, where $\chi^2$ is given by

$$\chi^2 = k\sigma_x^2/\sigma^2 \tag{4.19}$$

where $\sigma_x^2$ is the variance of a sample of $N$ measurements and $\sigma^2$ is the variance of the infinite data set that the sets of $N$ samples are part of. $k$ is a constant known as the number of degrees of freedom, and is equal to $(N - 1)$.

The shape of the chi-squared distribution depends on the value of $k$, with typical shapes being shown in Figure 4.5. The area under the chi-squared distribution curve is unity but, unlike the Gaussian distribution, the chi-squared distribution is not symmetrical. However, it tends toward the symmetrical shape of a Gaussian distribution as $k$ becomes very large.

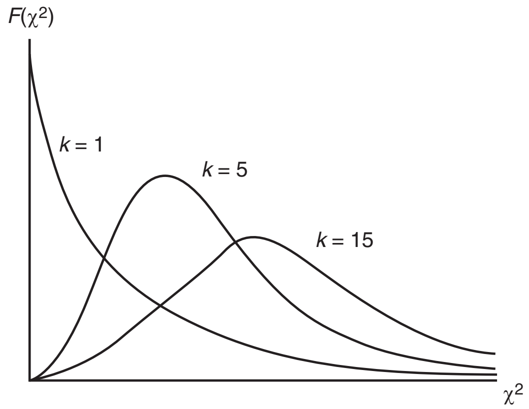


**Figure 4.5**
Typical chi-squared distributions.

The chi-squared distribution expresses the expected variation due to random chance of the variance of a sample away from the variance of the infinite population that the sample is part of. The magnitude of this expected variation depends on what level of "random chance" we set. The level of random chance is normally expressed as a *level of significance*, which is usually denoted by the symbol $\alpha$. Referring to the chi-squared distribution shown in Figure 4.6, the value $\chi^2_\alpha$ denotes the $\chi^2$ value to the left of which lies $100\,(1-\alpha)\%$ of the area under the $\chi^2$ distribution curve. Thus, the area of the curve to the right of $\chi^2_\alpha$ is $\alpha$ and that to the left is $(1-\alpha)$.

Numerical values for $\chi^2$ are obtained from tables that express the value of $\chi^2$ for various degrees of freedom $k$ and for various levels of significance $\alpha$. Published tables differ in the number of degrees of freedom and the number of levels of significance covered. A typical table is shown as Table 4.2. The first column in Table 4.2 gives various values of the

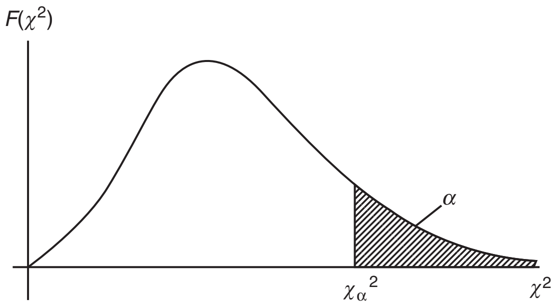


**Figure 4.6**
Meaning of symbol $\alpha$ for chi-squared distribution.

**Table 4.2: Chi-squared ($x^2$) distribution**

| $k$ | $\chi^2_{0.995}$ | $\chi^2_{0.990}$ | $\chi^2_{0.975}$ | $\chi^2_{0.950}$ | $\chi^2_{0.900}$ | $\chi^2_{0.100}$ | $\chi^2_{0.050}$ | $\chi^2_{0.025}$ | $\chi^2_{0.010}$ | $\chi^2_{0.005}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 2.71 | 3.84 | 5.02 | 6.63 | 7.88 |
| 2 | 0.01 | 0.02 | 0.05 | 0.10 | 0.21 | 4.61 | 5.99 | 7.38 | 9.21 | 10.6 |
| 3 | 0.07 | 0.12 | 0.22 | 0.35 | 0.58 | 6.25 | 7.81 | 9.35 | 11.3 | 12.8 |
| 4 | 0.21 | 0.30 | 0.48 | 0.71 | 1.06 | 7.78 | 9.49 | 11.1 | 13.3 | 14.9 |
| 5 | 0.41 | 0.55 | 0.83 | 1.15 | 1.61 | 9.24 | 11.1 | 12.8 | 15.1 | 16.7 |
| 6 | 0.68 | 0.87 | 1.24 | 1.64 | 2.20 | 10.6 | 12.6 | 14.4 | 16.8 | 18.5 |
| 7 | 0.99 | 1.24 | 1.69 | 2.17 | 2.83 | 12.0 | 14.1 | 16.0 | 18.5 | 20.3 |
| 8 | 1.34 | 1.65 | 2.18 | 2.73 | 3.49 | 13.4 | 15.5 | 17.5 | 20.1 | 22.0 |
| 9 | 1.73 | 2.09 | 2.70 | 3.33 | 4.17 | 14.7 | 16.9 | 19.0 | 21.7 | 23.6 |
| 10 | 2.16 | 2.56 | 3.25 | 3.94 | 4.87 | 16.0 | 18.3 | 20.5 | 23.2 | 25.2 |
| 11 | 2.60 | 3.05 | 3.82 | 4.57 | 5.58 | 17.3 | 19.7 | 21.9 | 24.7 | 26.8 |
| 12 | 3.07 | 3.57 | 4.40 | 5.23 | 6.30 | 18.5 | 21.0 | 23.3 | 26.2 | 28.3 |
| 13 | 3.57 | 4.11 | 5.01 | 5.89 | 7.04 | 19.8 | 22.4 | 24.7 | 27.7 | 29.8 |
| 14 | 4.07 | 4.66 | 5.63 | 6.57 | 7.79 | 21.1 | 23.7 | 26.1 | 29.1 | 31.3 |
| 15 | 4.60 | 5.23 | 6.26 | 7.26 | 8.55 | 22.3 | 25.0 | 27.5 | 30.6 | 32.8 |
| 16 | 5.14 | 5.81 | 6.91 | 7.96 | 9.31 | 23.5 | 26.3 | 28.8 | 32.0 | 34.3 |
| 17 | 5.70 | 6.41 | 7.56 | 8.67 | 10.1 | 24.8 | 27.6 | 30.2 | 33.4 | 35.7 |
| 18 | 6.26 | 7.02 | 8.23 | 9.39 | 10.9 | 26.0 | 28.9 | 31.5 | 34.8 | 37.2 |
| 19 | 6.84 | 7.63 | 8.91 | 10.1 | 11.7 | 27.2 | 30.1 | 32.9 | 36.2 | 38.6 |
| 20 | 7.43 | 8.26 | 9.59 | 10.9 | 12.4 | 28.4 | 31.4 | 34.2 | 37.6 | 40.0 |
| 21 | 8.03 | 8.90 | 10.3 | 11.6 | 13.2 | 29.6 | 32.7 | 35.5 | 38.9 | 41.4 |
| 22 | 8.64 | 9.54 | 11.0 | 12.3 | 14.0 | 30.8 | 33.9 | 36.8 | 40.3 | 42.8 |
| 23 | 9.26 | 10.2 | 11.7 | 13.1 | 14.9 | 32.0 | 35.2 | 38.1 | 41.6 | 44.2 |
| 24 | 9.89 | 10.9 | 12.4 | 13.8 | 15.7 | 33.2 | 36.4 | 39.4 | 43.0 | 45.6 |
| 25 | 10.5 | 11.5 | 13.1 | 14.6 | 16.5 | 34.4 | 37.7 | 40.6 | 44.3 | 46.9 |
| 26 | 11.2 | 12.2 | 13.8 | 15.4 | 17.3 | 35.6 | 38.9 | 41.9 | 45.6 | 48.3 |
| 27 | 11.8 | 12.9 | 14.6 | 16.2 | 18.1 | 36.7 | 40.1 | 43.2 | 47.0 | 49.6 |
| 28 | 12.5 | 13.6 | 15.3 | 16.9 | 18.9 | 37.9 | 41.3 | 44.5 | 48.3 | 51.0 |
| 29 | 13.1 | 14.3 | 16.0 | 17.7 | 19.8 | 39.1 | 42.6 | 45.7 | 49.6 | 52.3 |
| 30 | 13.8 | 15.0 | 16.8 | 18.5 | 20.6 | 40.3 | 43.8 | 47.0 | 50.9 | 53.7 |
| 35 | 17.2 | 18.5 | 20.6 | *12.5* | 24.8 | 46.1 | 49.8 | 53.2 | 57.3 | 60.3 |
| 40 | 20.7 | 22.2 | 24.4 | 26.5 | 29.1 | 51.8 | 55.8 | 59.3 | 63.7 | 66.8 |
| 45 | 24.3 | 25.9 | 28.4 | 30.6 | 33.4 | 57.5 | 61.7 | 65.4 | 70.0 | 73.2 |
| 50 | 28.0 | 29.7 | 32.4 | 34.8 | 37.7 | 63.2 | 67.5 | 71.4 | 76.2 | 79.5 |
| 75 | 47.2 | 49.5 | 52.9 | 56.1 | 59.8 | 91.1 | 96.2 | 100.8 | 106.4 | 110.3 |
| 100 | 67.3 | 70.1 | 74.2 | 77.9 | 82.4 | 118.5 | 124.3 | 129.6 | 135.8 | 140.2 |

degrees of freedom ($k$). The rest of the columns give the values of $\chi^2_\alpha$ for each value of $k$ corresponding to the values of $\alpha$ at the top of each column.

## ■ Example 4.7

Calculate $\chi^2$ for a sample of 10 measurements and a level of significance of 0.05.

■

## Solution

Referring to Table 4.2, the values of $\chi^2$ corresponding to a value $\alpha = 0.05$ are to be found in column 6 from the left of the table. For a sample size $N = 10$, the number of degrees of freedom $k$ is given by $(N - 1)$, that is, $k = 9$. Now reading along the row of the table corresponding to $k = 9$ as far as the 6th column gives a value for $\chi^2$ of 16.9.

## Example 4.8

Calculate $\chi^2$ for a sample of 21 measurements and a level of significance of 0.001.

## Solution

Referring to Table 4.2, the values of $\chi^2$ corresponding to a value $\alpha = 0.001$ are to be found in the final (rightmost) column of the table. For a sample size $N = 21$, the number of degrees of freedom $k$ is given by $(N - 1)$, that is, $k = 20$. Now reading along the row of the table corresponding to $k = 20$ as far as the last column gives a value for $\chi^2$ of 45.3.

One major use of the $\chi^2$ distribution is to predict the variance $\sigma^2$ of an infinite data set, given the measured variance $\sigma_x^2$ of a sample of $N$ measurements drawn from the infinite population. The boundaries of the range of $\chi^2$ values expected for a particular level of significance $\alpha$ can be expressed by the probability expression

$$P\left[\chi^2_{1-\alpha/2} \le \chi^2 \le \chi^2_{\alpha/2}\right] = 1 - \alpha \tag{4.20}$$

To put this in simpler terms, we are saying that there is a probability of $(1 - \alpha)\%$ that $\chi^2$ lies within the range bounded by $\chi^2_{1-\alpha/2}$ and $\chi^2_{\alpha/2}$ for a level of significance of $\alpha$. For example, for a level of significance $\alpha = 0.05$, there is a 95% probability (95% confidence level) that $\chi^2$ lies between $\chi^2_{0.975}$ and $\chi^2_{0.025}$.

Substituting into (4.20) using the expression for $\chi^2$ given in Eqn (4.19)

$$P\left[\chi^2_{1-\alpha/2} \le \frac{k\sigma_x^2}{\sigma^2} \le \chi^2_{\alpha/2}\right] = 1 - \alpha$$

This can be expressed in an alternative but equivalent form by inverting the terms and changing the "$\le$" relationships to "$\ge$" ones

$$P\left[\frac{1}{\chi^2_{1-\alpha/2}} \geq \frac{\sigma^2}{k\sigma_x^2} \geq \frac{1}{\chi^2_{\alpha/2}}\right] = 1 - \alpha$$

Now multiplying the expression through by $k_{\sigma_x^2}$ gives the following expression for the boundaries of the variance $\sigma^2$:

$$P\left[\frac{k\sigma_x^2}{\chi^2_{1-\alpha/2}} \geq \sigma^2 \geq \frac{k\sigma_x^2}{\chi^2_{\alpha/2}}\right] = 1 - \alpha \tag{4.21}$$

## Example 4.9

The length of each rod in a sample of 10 brass rods is measured and the variance of the length measurement in the sample is found to be 16.3 mm. Estimate the true variance and standard deviation for the whole batch of rods from which the sample of 10 was drawn, expressed to confidence level of 95%.

## Solution

Degrees of freedom $(k) = N - 1 = 9$.

For $\sigma_x^2 = 16.3, k\sigma_x^2 = 146.7$

For confidence level of 95%, level of significance, $\alpha = 0.05$.

Applying Eqn (4.21), the true variance is bounded by the values of $146.7/\chi^2_{0.975}$ and $146.7/\chi^2_{0.025}$.

Looking up the appropriate values in the $\chi^2$ distribution table for $k = 9$ gives

$$\chi^2_{0.975} = 2.70; \chi^2_{0.025} = 19.0; 146.7/\chi^2_{0.975} = 54.3; 146.7/\chi^2_{0.025} = 7.7$$

The true variance can therefore be expressed as $7.7 \leq \sigma^2 \leq 54.3$.

The true standard deviation can be expressed as $\sqrt{7.7} \leq \sigma \leq \sqrt{54.3}$,

that is, $2.8 \leq \sigma \leq 7.4$.

The thing that is immediately evident in this solution is that the range within which the true variance and standard deviation lie is very wide. This is a consequence of the relatively small number of measurements (10) in the sample. It is therefore highly desirable wherever possible to use a considerably larger sample when making predictions of the true variance and standard deviation of some measured quantity.

## Example 4.10

The length of a sample of 25 bricks is measured and the variance of the sample is calculated as 6.8 mm. Estimate the true variance for the whole batch of bricks from which the sample of 25 was drawn, expressed to confidence levels of (a) 90%, (b) 95%, and (c) 99%.

## Solution

Degrees of freedom $(k) = N - 1 = 24$.

For $\sigma_x^2 = 6.8, k\sigma_x^2 = 163.2$

(a) For confidence level of 90%, level of significance, $\alpha = 0.10$ and $\alpha/2 = 0.05$.

Applying Eqn (4.21), the true variance is bounded by the values of $163.2/\chi^2_{0.95}$ and $163.2/\chi^2_{0.05}$.

Looking up the appropriate values in the $\chi^2$ distribution table for $k = 24$ gives

$$\chi^2_{0.95} = 13.8;\;\; \chi^2_{0.05} = 36.4;\;\; 163.2/\chi^2_{0.95} = 11.8;\;\; 163.2/\chi^2_{0.05} = 4.5$$

The true variance can therefore be expressed as $4.5 \leq \sigma^2 \leq 11.8$.

(b) For confidence level of 95%, level of significance, $\alpha = 0.05$ and $\alpha/2 = 0.025$.

Applying Eqn (4.21), the true variance is bounded by the values of $163.2\,/\chi^2_{0.975}$ and $163.2/\chi^2_{0.025}$.

Looking up the appropriate values in the $\chi^2$ distribution table for $k = 24$ gives

$$\chi^2_{0.975} = 12.4;\;\; \chi^2_{0.025} = 39.4;\;\; 163.2/\chi^2_{0.975} = 13.2;\;\; 163.2/\chi^2_{0.025} = 4.1$$

The true variance can therefore be expressed as $4.1 \leq \sigma^2 \leq 13.2$.

(c) For confidence level of 99%, level of significance, $\alpha = 0.01$ and $\alpha/2 = 0.005$.

Applying Eqn (4.21), the true variance is bounded by the values of $146.7/\chi^2_{0.995}$ and $146.7/\chi^2_{0.005}$.

Looking up the appropriate values in the $\chi^2$ distribution table for $k = 24$ gives

$$\chi^2_{0.995} = 9.89;\;\; \chi^2_{0.005} = 45.6;\;\; 163.2/\chi^2_{0.995} = 16.5;\;\; 163.2/\chi^2_{0.005} = 3.6$$

The true variance can therefore be expressed as $3.6 \leq \sigma^2 \leq 16.5$.

The solution to Example (4.9) above shows that, as expected, the width of the estimated range in which the true value of standard deviation lies gets wider as we increase the confidence level from 90% to 99%. It is also interesting to compare the results in Examples (4.8) and (4.9) for the same confidence level of 95%. The ratio between the maximum and minimum values of estimated variance is much greater for the 10 samples in Example (4.8) compared with the 25 samples in Example (4.9). This shows the benefit of having a larger sample size when predicting the variance of the whole population that the sample is drawn from.

## 4.11 Goodness of Fit to a Gaussian Distribution

All of the analysis of random deviations presented so far only applies when the data being analyzed belong to a Gaussian distribution. Hence, the degree to which a set of data fits a Gaussian distribution should always be tested before any analysis is carried out. This test can be carried out in one of the three following ways:

### 4.11.1 Inspecting Shape of Histogram

The simplest way to test for Gaussian distribution of data is to plot a histogram and look for a "bell-shape" of the form shown earlier in Figure 4.1. Deciding whether or not the histogram confirms a Gaussian distribution is a matter of judgment. For a Gaussian distribution, there must always be approximate symmetry about the line through the center of the histogram, the highest point of the histogram must always coincide with this line of symmetry, and the histogram must get progressively smaller either side of this point. However, because the histogram can only be drawn with a finite set of measurements, some deviation from the perfect shape of histogram as described above is to be expected even if the data really are Gaussian.

Figure 4.7 shows example histograms of two sets of measurements. Histogram (a) is clearly not even close to being symmetrical about the vertical line through the center of the histogram, and shows that this data set is not Gaussian. In the case of histogram (b), there is approximate symmetry about the line through the center of the histogram. This indicates that the data probably follow a Gaussian distribution but further tests, by either drawing a normal probability test or applying the chi-squared test, are necessary to confirm whether the data set really is Gaussian.

### 4.11.2 Using a Normal Probability Plot

A normal probability plot involves dividing the data values into a number of ranges and plotting the cumulative probability of summed data frequencies against the data values on special graph paper.[1] This line should be a straight line if the data distribution is Gaussian.

[1] This is available from specialist stationery suppliers.

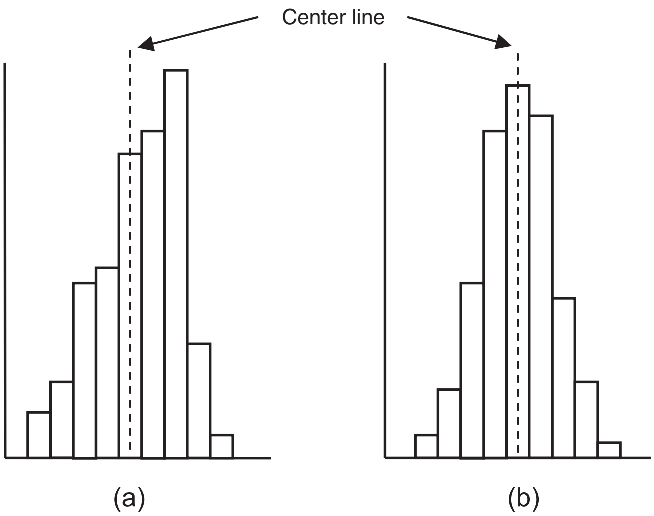


**Figure 4.7**
Histograms of data drawn to test goodness of fit to a Gaussian distribution. (a) Typical shape for non-Gaussian measurements; (b) typical shape for Gaussian measurements.

However, careful judgment is required since only a finite number of data values can be used and therefore the line drawn will not be entirely straight even if the distribution is Gaussian. Considerable experience is needed to judge whether the line is straight enough to indicate a Gaussian distribution. This will be easier to understand if the data in measurement set C is used as an example. Using the same five ranges as used to draw the histogram, the following table is first drawn:

| Range | 401.5–403.5 | 403.5–405.5 | 405.5–407.5 | 407.5–409.5 | 409.5–411.5 |
|---|---|---|---|---|---|
| Number of data items in range | 1 | 5 | 11 | 5 | 1 |
| Cumulative number of data items | 1 | 6 | 17 | 22 | 23 |
| Cumulative number of data items as % | 4.3 | 26.1 | 73.9 | 95.7 | 100.0 |

The normal probability plot drawn from the above table is shown in Figure 4.8. This is sufficiently straight to indicate that the data in measurement set C are Gaussian.

### 4.11.3 Chi-Squared Test

The chi-square distribution provides a more formal method for testing whether data follow a Gaussian distribution. The principle of the chi-squared test is to divide the data into $p$ equal width bins. The appropriate number for $p$ is usually determined by applying the Sturges. The next step is to count the number of measurements $n_i$ in each bin, using exactly the same procedure as done when drawing a histogram. The expected number of

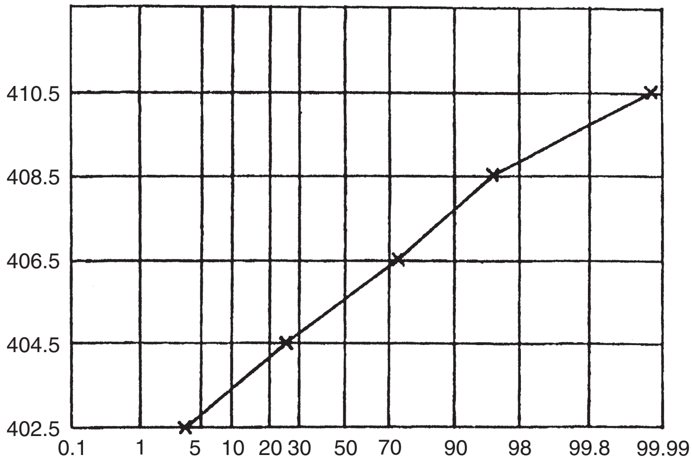


**Figure 4.8**
Normal probability plot.

measurements $n_i'$ in each bin for a Gaussian distribution is also calculated. Before proceeding any further, a check must be made at this stage to confirm that at least 80% of the bins have a data count greater than a minimum number for both $n_i$ and $n_i'$. We will apply a minimum number of 4, although some statisticians use the smaller minimum of 3 and some use a larger minimum of 5. If this check reveals that too many bins have data counts less than the minimum number, it is necessary to reduce the number of bins by redefining their widths. Common practice, when this is necessary, is to effect this reduction by halving the number of bins by combining them two at a time. In anticipation that this may be necessary, it is better to start initially with an even rather than odd number of bins. This means choosing the next highest even number if the Sturges rule suggests an odd number of bins. If reduction in the number of bins has been necessary, the test for at least 80% of the bins exceeding the minimum number then has to be reapplied. Once the data count in the bins is satisfactory, a chi-squared value is calculated for the data in each bin. Then, all the chi-squared values for the bins are added according to the following formula:

$$\chi^2 = \sum_{i=1}^{p} \frac{\left(n_i - n_i'\right)^2}{n_i'} \tag{4.22}$$

The chi-squared test then examines whether the calculated value of $\chi^2$ according to Eqn (4.22) is greater than would be expected for a Gaussian distribution according to some specified level of chance. This involves reading off the expected value from the chi-squared distribution table (Table 4.2) for the specified confidence level, and comparing this expected value with that calculated in Eqn (4.22). This procedure will become clearer if we work through an example.

## Example 4.11

A sample of 100 pork pies produced in a bakery is taken and the mass of each pie (grams) is measured. Apply the chi-squared test to examine whether the data set formed by the set of 100 mass measurements shown below conforms to a Gaussian distribution.

487 504 501 515 491 496 482 502 508 494 505 501 485 503 507 494 489 501 510 491
503 492 483 501 500 493 505 501 517 500 494 503 500 488 496 500 519 499 495 490
503 500 497 492 510 506 497 499 489 506 502 484 495 498 502 496 512 504 490 497
488 503 512 497 480 509 496 513 499 502 487 499 505 493 498 508 492 498 486 511
499 504 495 500 484 513 509 497 505 510 516 499 495 507 498 514 506 500 508 494

## Solution

Applying the Sturges rule, the recommended number of data bins $p$ for $N$ data points is given by

$p = 1 + 3.3 \log_{10} N = 1 + (3.3)(2.0000) = 7.6$. This rounds to 8.

The mass measurements span the range from 480 to 519. Hence we will choose data bin widths of 5 g, with bin boundaries set at 479.5, 484.5, 489.5, 494.5, 499.5, 504.5, 509.5, 514.5, 519.5 (boundaries are set so that there is no ambiguity about which bin any particular data value fits in). The next step involves counting the number of measurements in each bin. These are the $n_i$ values, $i = 1 \ldots 8$, for Eqn (4.22). The results of this counting are set out in the following table:

| Bin Number ($i$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Data range | 479.5 −484.5 | 484.5 −489.5 | 489.5 −494.5 | 494.5 −499.5 | 499.5 −504.5 | 504.5 −509.5 | 509.5 −514.5 | 514.5 −519.5 |
| Measurements in range ($n_i$) | 5 | 8 | 13 | 23 | 24 | 14 | 9 | 4 |

None of the bins have a count less than our stated minimum threshold of four and so we can now proceed to calculate the $n_i'$ values. These are the expected numbers of measurements in each data bin for a Gaussian distribution. The starting point for this calculation is knowing the mean value ($\mu$) and standard deviation ($\sigma$) of the 100 mass measurements. These are calculated using Eqns (4.1) and (4.7) as $\mu = 499.53$ and $\sigma = 8.389$. We now calculate the $z$ values corresponding to the measurement values ($x$) at the upper end of each data bin using Eqn (4.13) and then use the error function table (Table 4.1) to calculate $F(z)$. $F(z)$ gives the proportion of $z$ values that are $\leq z$, which gives the proportion of measurements less than the corresponding $x$ values. This then allows calculation of the expected number of measurements $(n_i')$ in each data bin. These calculations are shown in the following table:

| $x$ | $z\ \left(\frac{x-\mu}{\sigma}\right)$ | $F(z)$ | Expected Number of Data in Bin $(n_i')$ |
|---|---|---|---|
| 484.5 | −1.792 | 0.037 | 3.7 |
| 489.5 | −1.195 | 0.116 | 7.9 |
| 494.5 | −0.600 | 0.274 | 15.8 |
| 499.5 | −0.004 | 0.498 | 22.4 |
| 504.5 | 0.592 | 0.723 | 22.5 |
| 509.5 | 1.188 | 0.883 | 16.0 |
| 514.5 | 1.784 | 0.963 | 8.0 |
| 519.5 | 2.381 | 0.991 | 2.8 |

In case there is any confusion about the calculation of the numbers in the final column, let us consider rows 1 and 2. Row 1 shows that the proportion of data points less than 484.5 is 0.037. Since there are 100 data points in total, the actual estimated number of data points less than 484.5 is 3.7. Row 2 shows that the proportion of data points less than 489.5 is 0.116, and hence the total estimated number of data points less than 489.5 is 11.6. This total includes the 3.7 data points less than 484.5 calculated in the previous row. Hence, the number of data points in this bin between 484.5 and 489.5 is 11.6 − 3.7, that is, 7.9.

We can now calculate the $\chi^2$ value for the data using Eqn (4.22). The steps of the calculation are shown in the following table:

| Bin Number ($p$) | $n_i$ | $(n_i')$ | $(n_i - n_i')$ | $(n_i - n_i')^2$ | $\frac{(n_i - n_i')^2}{n_i'}$ |
|---|---|---|---|---|---|
| 1 | 5 | 3.7 | 1.3 | 1.69 | 0.46 |
| 2 | 8 | 7.9 | 0.1 | 0.01 | 0.00 |
| 3 | 13 | 15.8 | −2.8 | 7.84 | 0.50 |
| 4 | 23 | 22.4 | 0.6 | 0.36 | 0.02 |
| 5 | 24 | 22.5 | 1.5 | 2.25 | 0.10 |
| 6 | 14 | 16.0 | −2.0 | 4.00 | 0.25 |
| 7 | 9 | 8.0 | 1.0 | 1.00 | 0.12 |
| 8 | 4 | 2.8 | 1.2 | 1.44 | 0.51 |

The value of $\chi^2$ is now found by summing the values in the final column to give $\chi^2 = 1.96$. The final step is to check whether this value of $\chi^2$ is greater than would be expected for a Gaussian distribution. This involves looking up $\chi^2$ in Table 4.2. Before doing this, we have to specify the number of degrees of freedom, $k$. In this case, $k$ is the number of bins minus 2, because the data are manipulated twice to obtain the $\mu$ and $\sigma$ statistical values used in the calculation of $n_i'$. Hence, $k = 8-2 = 6$.

Table 4.2 shows that, for $k = 6$, $\chi^2 = 1.64$ for a 95% confidence level and $\chi^2 = 2.20$ for a 90% confidence level. Hence, our calculated value for $\chi^2$ of 1.96 shows that the confidence level that the data follow a Gaussian distribution is between 90% and 95%.

■

We will now look at a slightly different example, where we meet the problem that our initial division of the data into bins produces too many bins that do not contain the minimum number of data points necessary for the chi-squared test to work reliably.

## ■ Example 4.12

Suppose that the production machinery used to produce the pork pies featured in Example (4.11) is modified to try and reduce the amount of variation in mass. The mass of a new sample of 100 pork pies is then measured. Apply the chi-squared test to examine whether the data set formed by the set of 100 new mass measurements shown below conforms to a Gaussian distribution.

503 509 495 500 504 491 496 499 501 489 507 501 486 497 500 493 499 505 501 495
499 515 505 492 499 502 507 500 498 507 494 499 506 501 493 498 505 499 496 512
498 502 508 500 497 485 504 499 502 496 483 501 510 494 498 505 491 499 503 495
502 481 498 503 508 497 511 490 506 500 508 504 517 494 487 505 499 509 492 484
500 507 501 496 510 503 498 490 501 492 497 489 502 495 491 500 513 499 494 498

■

## ■ Solution

The recommended number of data bins for 100 measurements according to the Sturges rule is 8, as calculated in Example (4.11). The mass measurements in this new data set span the range from 481 to 517. Hence, data bin widths of 5 g are still suggested, with bin boundaries set at 479.5, 484.5, 489.5, 494.5, 499.5, 504.5, 509.5, 514.5, 519.5. The number of measurements in each bin is then counted, with the counts given in the following table:

| Bin number ($i$) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Data range | 479.5 −484.5 | 484.5 −489.5 | 489.5 −494.5 | 494.5 −499.5 | 499.5 −504.5 | 504.5 −509.5 | 509.5 −514.5 | 514.5 −519.5 |
| Measurements in range $n_i$ | 3 | 5 | 14 | 29 | 26 | 16 | 5 | 2 |

Looking at these counts, we see that there are two bins with a count less than 4. This amounts to 25% of the data bins. We have previously said that not more than 20% of the data bins can have a data count less than the threshold of 4 if the chi-squared test is to operate reliably. Hence, we must combine the bins and count the measurements again. The usual approach is to combine pairs of bins, which in this case reduces the number of bins from eight to four. The boundaries of the new set of four bins are now 479.5, 489.5, 499.5, 509.5 and 519.5. The new data ranges and counts are shown in the following table:

| Bin number ($i$) | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Data range | 479.5−489.5 | 489.5−499.5 | 499.5−509.5 | 509.5−519.5 |
| Measurements in range $n_i$ | 8 | 43 | 42 | 7 |

Now, none of the bins have a count less than our stated minimum threshold of four and so we can proceed to calculate the $n_i'$ values as before. The mean value ($\mu$) and standard deviation ($\sigma$) of the new mass measurements are $\mu = 499.39$ and $\sigma = 6.979$. We now calculate the $z$ values corresponding to the measurement values ($x$) at the upper end of each data bin, read off the corresponding $F(z)$ values from Table 4.1 and so calculate the expected number of measurements $(n_i')$ in each data bin:

| $x$ | $z$ $\left(\frac{x-\mu}{\sigma}\right)$ | $F(z)$ | Expected Number of Data in Bin $(n_i')$ |
|---|---|---|---|
| 489.5 | −1.417 | 0.078 | 7.8 |
| 499.5 | −0.016 | 0.494 | 41.6 |
| 509.5 | 1.449 | 0.926 | 43.2 |
| 519.5 | 2.882 | 0.998 | 7.2 |

Note: Since the values for $z$ are calculated to an accuracy of three figures after the decimal point but the error function table used to calculate $F(z)$ only gives values for $z$ with a maximum of two figures after the decimal point, interpolation between the nearest $F(z)$ values has to be carried out, the details of which are explained in Appendix 4

In case there is any confusion about the calculation of the numbers in the final column, let us consider rows 1 and 2. Row 1 shows that the proportion of data points less than 489.5 is 0.078. Since there are 100 data points in total, the actual estimated number of data points less than 484.5 is 7.8. Row 2 shows that the proportion of data points less than 499.5 is 0.494, and hence the total estimated number of data points less than 499.5 is 49.4. This total includes the 7.8 data points less than 489.5 calculated in the previous row. Hence, the number of data points in this bin between 489.5 and 499.5 is 49.4 − 7.8, that is, 41.6.

We now calculate the $\chi^2$ value for the data using Eqn (4.22). The steps of the calculation are shown in the table below:

| Bin Number ($p$) | $n_i$ | $(n_i')$ | $(n_i - n_i')$ | $(n_i - n_i')^2$ | $\frac{(n_i - n_i')^2}{n_i'}$ |
|---|---|---|---|---|---|
| 1 | 8 | 7.8 | 0.2 | 0.04 | 0.005 |
| 2 | 43 | 41.6 | 1.4 | 1.96 | 0.047 |
| 3 | 42 | 43.2 | −1.2 | 1.44 | 0.033 |
| 4 | 7 | 7.2 | −0.2 | 0.04 | 0.006 |

The value of $\chi^2$ is now found by summing the values in the final column to give $\chi^2 = 0.091$. The final step is to check whether this value of $\chi^2$ is greater than would be expected for a Gaussian distribution. This involves looking up $\chi^2$ in Table 4.2. This time, $k = 2$, since there are 4 bins and $k$ is the number of bins minus 2 (as explained in Example 4.11, the data were manipulated twice to obtain the $\mu$ and $\sigma$ statistical values used in the calculation of $(n_i')$).

Table 4.2 shows that, for $k = 2$, $\chi^2 = 0.10$ for a 95% confidence level. Hence, our calculated value for $\chi^2$ of 0.91 shows that the confidence level that the data follow a Gaussian distribution is slightly better than 95%.

■

Out of interest, if the 2 bin counts less than 4 had been ignored and $\chi^2$ had been calculated for the 8 original data bins, a value of $\chi^2 = 2.97$ would have been obtained. (It would be a useful exercise for the reader to check this for himself/herself.) For six degrees of freedom ($k = 8 - 2$), the predicted value of $\chi^2$ for a Gaussian population from Table 4.2 is 2.20 at a 90% confidence level. Thus the confidence that the data fit a Gaussian distribution is substantially less than 90% given the $\chi^2$ value of 2.97 calculated for the data. This result arises because of the unreliability associated with calculating $\chi^2$ from data bin counts of less than 4.

## 4.12 Rogue Data Points (Data Outliers)

In a set of measurements subject to random error, measurements with a very large error sometimes occur at random and unpredictable times, where the magnitude of the error is much larger than could reasonably be attributed to the expected random variations in measurement value. These are often called *rogue data points* or *data outliers*. Sources of such abnormal error include sudden transient voltage surges on the mains power supply and incorrect recording of data (e.g., writing down 146.1 when the actual measured value was 164.1). It is accepted practice in such cases to discard these rogue measurements, and a threshold level of $\pm 3\sigma$ deviation is often used to determine what should be discarded. It is rare for measurement errors to exceed $\pm 3\sigma$ limits when only normal random effects are affecting the measured value.

While the above represents a reasonable theoretical approach to identifying and eliminating rogue data points, the practical implementation of such a procedure needs to be done with care. The main practical difficulty that exists in dealing with rogue data points is in establishing what the expected standard deviation of the measurements is. When a new set of measurements is being taken where the expected standard deviation is not known, the possibility exists that a rogue data point exists within the measurements. Simply applying a computer program to the measurements to calculate the standard deviation will produce an erroneous result because the calculated value will be biased by the rogue data point. The simplest way to overcome this difficulty is to plot a histogram of any new set of measurements and examine this manually to spot any data outliers. If no outliers are apparent, the standard deviation can be calculated and then used in a $\pm 3\sigma$ threshold against which to test all future measurements. However, if this initial data histogram shows up any outliers, these should be excluded from the calculation of the standard deviation.

### Example 4.13

A set of measurements is made with a new pressure transducer. Inspection of a histogram of the first 20 measurements does not show any data outliers. The standard deviation of the measurements is calculated as 0.05 bar after this check for data outliers, and the mean value is calculated as 4.41. Following this, further set of measurements is obtained:

4.35 4.46 4.39 4.34 4.41 4.52 4.44 4.37 4.41 4.33 4.39 4.47 4.42 4.59 4.45 4.38 4.43 4.36 4.48 4.45

Use the $\pm 3\sigma$ threshold to determine whether there are any rogue data points in the measurement set.

■

## Solution

Since the calculated $\sigma$ value for a set of "good" measurements is given as 0.05, the $\pm 3\sigma$ threshold is $\pm 0.15$. With a mean data value of 4.41, the threshold for rogue data points is values below 4.26 (mean value minus $3\sigma$) or above 4.56 (mean value plus $3\sigma$). Looking at the set of measurements, we observe that the measurement of 4.59 is outside the $\pm 3\sigma$ threshold, indicating that this is a rogue data point.

It is interesting at this point to return to the problem of ensuring that there are no outliers in the set of data used to calculate the standard deviation of the data and hence the threshold for rejecting outliers. We have suggested that a histogram of some initial measurements be drawn and examined for outliers. What would happen if the set of data given above in Example 4.13 was the initial data set that was examined for outliers by drawing a histogram? What would happen if we did not spot the outlier of 4.59? This question can be answered by looking at the effect on the calculated value of standard deviation if this rogue data point of 4.59 is included in the calculation. The standard deviation calculated over the 19 values excluding the 4.59 measurement is 0.052. The standard deviation calculated over the 20 values including the 4.59 measurement is 0.063 and the mean data value is changed to 4.42. This gives a $3\sigma$ threshold of 0.19, and the boundaries for the $\pm 3\sigma$ threshold operation are now 4.23 and 4.61. This does not exclude the data value of 4.59 which we previously identified as being a rogue data point! This confirms the necessity of carefully looking at the initial set of data used to calculate the thresholds for rejection of rogue data point to ensure that the initial data do not contain any rogue data points. If drawing and examining a histogram do not clearly show that there are no rogue data points in the "reference" set of data, it is worth taking another set of measurements to see whether a reference set of data can be obtained that is more clearly free of rogue data points.

The discussion so far has suggested that it is always necessary to have a reference data set, where the absence of rogue data points has been confirmed by drawing a histogram. However, there are some occasions, where a rogue data point can be spotted even with the tedium of drawing a histogram, as demonstrated by the example below.

## Example 4.14

Consider the following data set of 20 measurements and examine it for rogue data points:

10.4 9.9 9.7 9.6 10.1 10.3 9.8 10.0 10.2 9.5 9.8 10.1 10.3 8.1 9.7 10.2 10.3 9.7 9.9 10.2

## Solution

The measurement of 8.1 clearly looks to be a rogue data point. However, this must still be confirmed. The mean and standard deviation of the other 19 data points (excluding the value of 8.1) are calculated as mean = 9.98 and standard deviation ($\sigma$) = 0.273. This means that $3\sigma = 0.819$. Hence, the $\pm 3\sigma$ limits either side of the mean value are at 9.16 and 10.80. The data point at 8.1 is clearly well outside the $\pm 3\sigma$ limits and it is therefore confirmed as a rogue data point.

## 4.13 Student t-Distribution

When the number of measurements of a quantity is particularly small (less than about 30 samples) and statistical analysis of the distribution of error values is required, the possible deviation of the mean of the measurements from the true measurement value (the mean of the infinite population that the sample is part of) may be significantly greater than is suggested by analysis based on a $z$-distribution. In response to this, a statistician called William Gosset developed an alternative distribution function that gives a more accurate prediction of the error distribution when the number of samples is small. He published this under the pseudonym "Student" and the distribution is commonly called the *Student t-distribution*. It should be noted that the $t$-distribution has the same requirement as the $z$-distribution in terms of the necessity for the data to belong to a Gaussian distribution.

The Student $t$-variable expresses the difference between the mean of a small sample ($x_{\text{mean}}$) and the population mean ($\mu$) in terms of the following ratio:

$$t = \frac{|\text{error in mean}|}{\text{standard error of the mean}} = \frac{|\mu - x_{\text{mean}}|}{\sigma / \sqrt{N}} \tag{4.23}$$

where $N$ is the number of samples, $\sigma$ is the standard error of the mean of the infinite population that the $N$ samples are part of. Since we do not know the exact value of $\sigma$, we have to use the best approximation to $\sigma$ that we have, which is the standard deviation of the sample $\sigma_x$. Substituting this value for $\sigma$ in (4.23) gives

$$t = \frac{|\mu - x_{\text{mean}}|}{\sigma_x / \sqrt{N}} \tag{4.24}$$

Note that the modulus operation ($|\cdots|$) on the error in the mean in Eqns (4.23) and (4.24) means that $t$ is always positive.

The shape of the probability distribution curve $F(t)$ of the $t$-variable varies according to the value of the number of degrees of freedom, $k$ ($= N - 1$), with typical curves being shown in Figure 4.9. As $k \to \infty$, $F(t) \to F(z)$, that is, the distribution becomes a standard Gaussian one. For values of $k < \infty$, the curve of $F(t)$ against $t$ is both narrower and less high in the center than a standard Gaussian curve, but has the same properties of symmetry about $t = 0$ and a total area under the curve of unity.

In a similar way to the $z$-distribution, the probability that $t$ will lie between two values $t_1$ and $t_2$ is given by the area under the $F(t)$ curve between $t_1$ and $t_2$. The $t$-distribution is published in the form of a standard table (see Table 4.3) that gives values of the area under the curve $\alpha$ for various values of the number of degrees of freedom ($k$), where

$$\alpha = \int_{t_\alpha}^{\infty} F(t)\mathrm{d}t \tag{4.25}$$

The area $\alpha$ is shown in Figure 4.10. $\alpha$ corresponds to the probability that $t$ will have a value greater than $t_\alpha$ to some specified confidence level. Because the total area under the $F(t)$ curve is unity, there is also a probability of $(1 - \alpha)$ that $t$ will have a value less than $t_\alpha$. Thus, for a value $\alpha = 0.05$, there is a 95% probability (i.e., a 95% confidence level) that $t < t_\alpha$.

Because of the symmetry of the $t$-distribution, $\alpha$ is also given by

$$\alpha = \int_{-\infty}^{-t_\alpha} F(t)\mathrm{d}t \tag{4.26}$$

as shown in Figure 4.11. Here, $\alpha$ corresponds to the probability that $t$ will have a value less than $-t_\alpha$, with a probability of $(1 - \alpha)$ that $t$ will have a value greater than $-t_\alpha$.

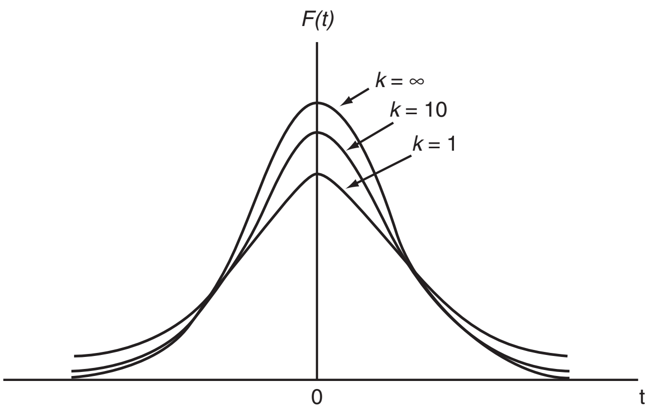


**Figure 4.9**
Typical $t$-distribution curves.

**Table 4.3: *t*-distribution**

| $k$ | $t_{0.10}$ | $t_{0.05}$ | $t_{0.025}$ | $t_{0.01}$ | $t_{0.005}$ | $t_{0.001}$ |
|---|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.71 | 31.82 | 63.66 | 318.3 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 23.33 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.21 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 3.610 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 3.527 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 3.505 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 3.485 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 3.467 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 3.450 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 3.435 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 3.421 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 3.408 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 3.396 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.385 |

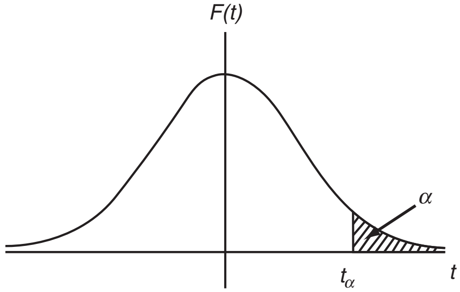


**Figure 4.10**
Meaning of area $\alpha$ for $t$-distribution curve.

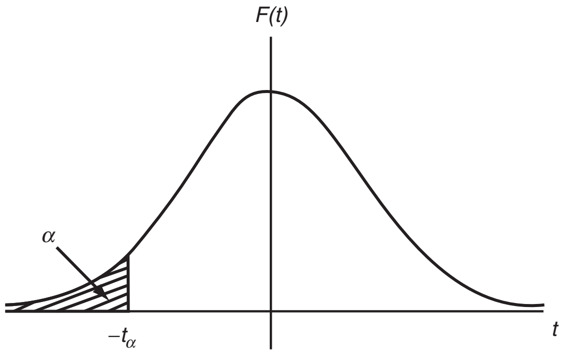


**Figure 4.11**
Alternative interpretation of area $\alpha$ for $t$-distribution curve.

## Example 4.15

Calculate the value of the variable $t$ for a set of 15 measurements and a 97.5% confidence level.

## Solution

The number of degrees of freedom ($k$) is one less than the number of measurements. Thus $k = 14$.

For a 97.5% confidence level, $\alpha = 1 - 0.975 = 0.025$.

The values of $t$ corresponding to a value $\alpha = 0.025$ are to be found in the fourth column of Table 4.3. Reading along the row in the table corresponding to $k = 14$ as far as the 4th column gives a value of $t$ of 2.145.

Equations (4.25) and (4.26) can be combined to express the probability $(1 - \alpha)$ that $t$ lies between two values $-t_4$ and $+t_4$. In this case, $\alpha$ is the sum of two areas of $\alpha/2$ as shown in Figure 4.12. These two areas can be represented mathematically as

$$\frac{\alpha}{2} = \int_{-\infty}^{-t_4} F(t)\mathrm{d}t \quad \text{(left-hand area)} \qquad \text{and} \qquad \frac{\alpha}{2} = \int_{t_4}^{\infty} F(t)\mathrm{d}t \quad \text{(right-hand area)}$$

The values of $t_4$ can be found any $t$-distribution table, such as Table 4.3.

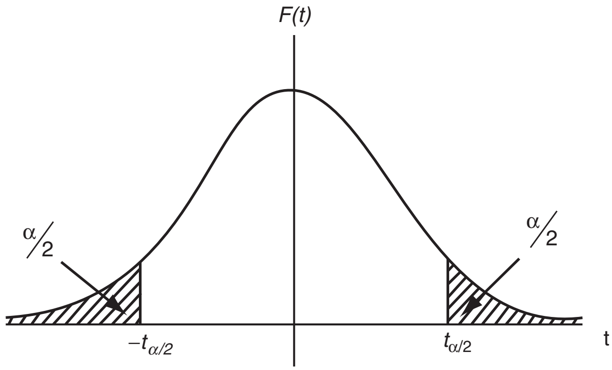


**Figure 4.12**
Area between $-\alpha/2$ and $+\alpha/2$ on $t$-distribution curve.

Referring back to Eqn (4.24), this can be expressed in the form

$$\left|\mu - x_{\text{mean}}\right| = \frac{t\sigma_x}{\sqrt{N}}$$

Hence, the upper and lower bounds on the expected value of the population mean $\mu$ (the true value of $x$) can be expressed as

$$-\frac{t_4\sigma_x}{\sqrt{N}} \le \mu - x_{\text{mean}} \le +\frac{t_4\sigma_x}{\sqrt{N}}$$

or

$$x_{\text{mean}} - \frac{t_4\sigma_x}{\sqrt{N}} \le \mu \le x_{\text{mean}} + \frac{t_4\sigma_x}{\sqrt{N}} \tag{4.27}$$

## Example 4.16

The internal diameter of a sample of hollow castings is measured by destructive testing of 15 samples taken randomly from a large batch of castings. If the sample mean is 105.4 mm with a standard deviation of 1.9 mm, express the upper and lower bounds to a confidence level of 95% on the range in which the mean value lies for internal diameter of the whole batch.

## Solution

For 15 samples ($N = 15$), the number of degrees of freedom ($k$) = 14.

For a confidence level of 95%, $\alpha = 1 - 0.95 = 0.05$. Looking up the value of $t$ in Table 4.3 for $k = 14$ and $\alpha/2 = 0.025$ gives $t = 2.145$. Thus, applying Eqn (4.27)

$$105.4 - \frac{(2.145)(1.9)}{\sqrt{15}} \leq \mu \leq 105.4 + \frac{(2.145)(1.9)}{\sqrt{15}},$$

that is, $104.3 \leq \mu \leq 106.5$.
Thus, we would express the mean internal diameter of the whole batch of castings as $105.4 \pm 1.1$ mm.

## Example 4.17

The width of the air gap in some double glazing panels is measured by destructive testing of six panels. If the mean gap for the sample is 19.7 mm with a standard deviation of 0.6 mm, express the upper and lower bounds to a confidence level of 99% on the range in which the mean value lies for the air gap in the whole batch.

## Solution

For six samples ($N = 6$), the number of degrees of freedom ($k$) = 5.

For a confidence level of 99%, $\alpha = 1 - 0.99 = 0.01$. Looking up the value of $t$ in Table 4.3 for $k = 5$ and $\alpha/2 = 0.005$ gives $t = 4.032$. Thus, applying Eqn (4.27)

$$19.7 - \frac{(4.032)(0.6)}{\sqrt{6}} \leq \mu \leq 19.7 + \frac{(4.032)(0.6)}{\sqrt{6}},$$

that is, $18.7 \leq \mu \leq 20.7$.
Thus, we would express the mean internal diameter of the whole batch of panels as $19.7 \pm 1.0$ mm.

Out of interest, let us examine what would have happen if we had calculated the error bounds on $\mu$ using standard Gaussian ($z$-distribution) tables. For 95% confidence, the maximum error is given as $\pm 1.96\sigma/\sqrt{N}$, that is, $\pm 0.96$ which rounds to $\pm 1.0$ mm, meaning the mean internal diameter is given as $105.4 \pm 1.0$ mm. The effect of using the $t$-distribution instead of the $z$-distribution clearly expands the magnitude of the likely error in the mean value to compensate for the fact that our calculations are based on a relatively small number of measurements.

## 4.14 Aggregation of Measurement System Errors

Errors in measurement systems often arise from two or more different sources, and these must be aggregated in the correct way in order to obtain a prediction of the total likely error in output readings from the measurement system. Two different forms of aggregation are required. First, a single measurement component may have both systematic and random errors and, second, a measurement system may consist of several measurement components that each have separate errors.

### 4.14.1 Combined Effect of Systematic and Random Errors

If a measurement is affected by both systematic and random errors that are quantified as $\pm x$ (systematic errors) and $\pm y$ (random errors), some means of expressing the combined effect of both types of error are needed. One way of expressing the combined error would be to sum the two separate components of error, that is, to say that the total possible error is $e = \pm(x + y)$. However, a more usual course of action is to express the likely maximum error as

$$e = \sqrt{(x^2 + y^2)} \tag{4.28}$$

It can be shown (ANSI/ASME, 1998) that this is the best expression for the error statistically, since it takes account of the reasonable assumption that the systematic and random errors are independent and so it is unlikely to both will be at their maximum or minimum value simultaneously.

### 4.14.2 Aggregation of Errors from Separate Measurement System Components

A measurement system often consists of several separate components, each of which is subject to errors. Therefore, what remains to be investigated is how the errors associated with each measurement system component combine together, so that a total error calculation can be made for the complete measurement system. All four mathematical operations of addition, subtraction, multiplication, and division may be performed on measurements derived from different instruments/transducers in a measurement system. Appropriate techniques for the various situations that arise are covered below.

*Error in a sum*

If the two outputs $y$ and $z$ of separate measurement system components are to be added together, we can write the sum as $S = y + z$. If the maximum errors in $y$ and $z$ are $\pm ay$ and $\pm bz$ respectively, we can express the maximum and minimum possible values of $S$ as

$$S_{\max} = (y + ay) + (z + bz); \quad S_{\min} = (y - ay) + (z - bz); \quad \text{or} \quad S = y + z \pm (ay + bz)$$

This relationship for $S$ is not convenient, because in this form the error term cannot be expressed as a fraction or percentage of the calculated value for $S$. Fortunately, statistical analysis can be applied (see Topping, 1972 for details.) that expresses $S$ in an alternative form such that the most probable maximum error in $S$ is represented by a quantity $e$, where $e$ is calculated in terms of the *absolute* errors as

$$e = \sqrt{(ay)^2 + (bz)^2} \tag{4.29}$$

Thus $S = (y + z) \pm e$. This can be expressed in the alternative form

$$S = (y + z)(1 \pm f) \tag{4.30}$$

where $f = e/(y + z)$.

It should be noted that Eqns (4.29) and (4.30) are only valid provided that the measurements are uncorrelated (i.e., each measurement is entirely independent of the others).

## Example 4.18

A circuit requirement for a resistance of 550Ω is satisfied by connecting together two resistors of nominal values 220 and 330 Ω in series. If each resistor has a tolerance of ±2%, the error in the sum calculated according to Eqns (4.29) and (4.30) is given by

$$e = \sqrt{(0.02 \times 220)^2 + (0.02 \times 330)^2} = 7.93; \quad f = 7.93/50 = 0.0144$$

Thus the total resistance $S$ can be expressed as $S = 550\ \Omega \pm 7.93\ \Omega$ or $S = 550\ (1 \pm 0.0144)\ \Omega$, that is, $S = 550\ \Omega \pm 1.4\%$.

*Error in a difference*

If the two outputs $y$ and $z$ of separate measurement systems are to be subtracted from one another, and the possible errors are $\pm ay$ and $\pm bz$, then the difference $S$ can be expressed (using statistical analysis as for calculating the error in a sum and assuming that the measurements are uncorrelated) as

$$S = (y - z) \pm e \quad \text{or} \quad S = (y - z)(1 \pm f)$$

where $e$ is calculated as above (Eqn (4.29)), and $f = e/(y - z)$.

## Example 4.19

A fluid flow rate is calculated from the difference in pressure measured on both sides of an orifice plate. If the pressure measurements are 10.0 bar and 9.5 bar and the error in the pressure-measuring instruments is specified as ±0.1%, then values for $e$ and $f$ can be calculated as

$$e = \sqrt{(0.001 \times 10)^2 + (0.001 \times 9.5)^2} = 0.0138; \;\; f = 0.0138/0.5 = 0.0276$$

■

- *This example illustrates very poignantly the relatively large error that can arise when calculations are made based on the difference between two measurements.*

*Error in a product*

If the outputs $y$ and $z$ of two measurement system components are multiplied together, the product can be written as $P = yz$. If the possible error in $y$ is $\pm ay$ and in $z$ is $\pm bz$, then the maximum and minimum values possible in $P$ can be written as

$$P_{\max} = (y + ay)(z + bz) = yz + ayz + byz + aybz;$$
$$P_{\min} = (y - ay)(z - bz) = yz - ayz - byz + aybz$$

For typical measurement system components with output errors of up to 1 or 2% in magnitude, both $a$ and $b$ are very much less than one in magnitude and thus terms in $aybz$ are negligible compared with other terms. Therefore, we have $P_{\max} = yz(1 + a + b)$; $P_{\min} = yz(1 - a - b)$. Thus the maximum error in the product $P$ is $\pm(a + b)$. While this expresses the maximum possible error in $P$, it tends to overestimate the likely maximum error since it is very unlikely that the errors in $y$ and $z$ will both be at the maximum or minimum value at the same time. A statistically better estimate of the likely maximum error $e$ in the product $P$, provided that the measurements are uncorrelated, is given by (for derivation, see Topping, 1972).

$$e = \sqrt{a^2 + b^2} \tag{4.31}$$

Note that in the case of multiplicative errors, $e$ is calculated in terms of the *fractional* errors in $y$ and $z$ (as opposed to the *absolute* error values used in calculating additive errors).

## ■ Example 4.20

If the power in a circuit is calculated from measurements of voltage and current in which the calculated maximum errors are respectively ±1% and ±2%, then the maximum likely error in the power value, calculated using (4.31) is $\pm\sqrt{0.01^2 + 0.02^2} = \pm 0.022$ or $\pm 2.2\%$.

■

*Error in a quotient*

If the output measurement $y$ of one system component with possible error $\pm ay$ is divided by the output measurement $z$ of another system component with possible error $\pm bz$, then the maximum and minimum possible values for the quotient can be written as

$$Q_{\max} = \frac{y+ay}{z-bz} = \frac{(y+ay)(z+bz)}{(z-bz)(z+bz)} = \frac{yz+ayz+byz+aybz}{z^2-b^2z^2};$$

$$Q_{\min} = \frac{y-ay}{z+bz} = \frac{(y-ay)(z-bz)}{(z+bz)(z-bz)} = \frac{yz-ayz-byz+aybz}{z^2-b^2z^2}$$

For $a \ll 1$ and $b \ll 1$, terms in $ab$ and $b^2$ are negligible compared with the other terms.

Hence $Q_{\max} = \frac{yz(1+a+b)}{z^2}$; $Q_{\min} = \frac{yz(1-a-b)}{z^2}$; that is $Q = \frac{y}{z} \pm \frac{y}{z}(a+b)$.

Thus the maximum error in the quotient is $\pm(a+b)$. However, using the same argument as made above for the product of measurements, a statistically better estimate of the likely maximum error in the quotient $Q$, provided that the measurements are uncorrelated, is that given in (4.31).

## Example 4.21

If the density of a substance is calculated from measurements of its mass and volume where the respective errors are $\pm 2\%$ and $\pm 3\%$, then the maximum likely error in the density value using (4.31) is $\pm\sqrt{0.02^2+0.03^2} = \pm 0.036$ or $\pm 3.6\%$.

### 4.14.3 Total Error When Combining Multiple Measurements

The final case to be covered is where the final measurement is calculated from several measurements that are combined together in a way that involves more than one type of arithmetic operation. For example, the density of a rectangular-sided solid block of material can be calculated from measurements of its mass divided by the product of measurements of its length, height, and width. The errors involved in each stage of arithmetic are cumulative, and so the total measurement error can be calculated by adding together the two error values associated with the two multiplication stages involved in calculating the volume and then calculating the error in the final arithmetic operation when the mass is divided by the volume.

## Example 4.22

A rectangular-sided block has edges of lengths $a$, $b$, and $c$, and its mass is $m$. If the values and possible errors in quantities $a$, $b$, $c$, and $m$ are as shown below, calculate the value of density and the possible error in this value.

$$a = 100\ \text{mm} \pm 1\%,\quad b = 200\ \text{mm} \pm 1\%,\quad c = 300\ \text{mm} \pm 1\%,\quad m = 20\ \text{Kg} \pm 0.5\%.$$

## Solution

$$\text{Value of } ab = 0.02\ \text{m}^2 \pm 2\% \text{ (possible error} = 1\% + 1\% = 2\%)$$

$$\text{Value of } (ab)c = 0.006\ \text{m}^3 \pm 3\% \text{ (possible error} = 2\% + 1\% = 3\%)$$

$$\text{Value of } \frac{m}{abc} = \frac{20}{0.006} = 3330\ \text{kg/m}^3 \pm 3.5\% \text{ (possible error} = 3\% + 0.5\% = 3.5\%)$$

## *4.15 Summary*

This chapter has been the second of two chapters dealing with measurement uncertainty. Chapter 3 introduced the subject and established that there were two distinct forms of measurement uncertainty, known respectively as *systematic error* and *random error*. This last chapter then went on to consider systematic errors in detail, and in particular, the sources of such errors, ways of reducing their magnitude, and the techniques for quantifying the remaining error after all reasonable means of reducing error magnitude had been applied. This chapter then went on to introduce random errors and explain their nature and typical sources. There was also a brief introduction to the means available for quantifying random errors. This current chapter has continued, where Chapter 3 left off in examining random errors in much greater detail and describing the various means available for analyzing random errors.

The starting point in this chapter was to examine the two alternative ways of calculating the average value of a set of measurements of a constant quantity, these being the *mean* and *median* values. We then went on to explain how the way in which measurements are spread about the average value affects the degree of confidence that the calculated mean value is close to the correct value of the measured quantity. We explained the calculation of the two quantities *standard deviation* and *variance* of measurement data, these being parameters that express how the measurements are distributed about the mean value.

Following on from this, we started to look at graphical ways of expressing the spread. Initially, we considered representations of spread as a *histogram*, and then went on to show how histograms expand into *frequency distributions* in the form of a smooth curve. We found that truly random data are described by a particular form of frequency distribution known as *Gaussian* (or *normal*). We introduced the *z variable* and saw how this can be used to estimate the number of measurements in a set of measurements that have an error magnitude between two specified values. Following this, we started to look at the implications of the fact that we can only ever have a finite number of measurements. We saw that a variable called the *standard error of the mean* could be calculated, which estimates the difference between the mean value of a finite set of measurements and the true value of the measured quantity (the mean of an infinite data set). We went on to look at how this was useful in estimating the likely error in a single measurement subject to random errors, in the situation, where it is not possible to average over a number of measurements. As an aside, we then went on to look at how the $z$ variable was useful in analyzing the tolerances of manufactured components subject to random variations, in a parallel way to the analysis of measurements subject to random variations. Following this, we went on to look at the chi-squared distribution. This can be used to quantify the variation in the variance of a finite set of measurements with respect to the variance of the infinite set that the finite set is part of. Up to this point in the chapter, all analysis of random errors assumed that the measurement set fitted a Gaussian distribution. However, this assumption must always be justified by applying *goodness of fit tests*, and so these were explained in the following section, where we saw that a chi-squared test is the most rigorous test available for goodness of fit. A particular problem that can adversely affect the analysis of random errors is the presence of rogue data points (data outliers) in the measurement data. These were considered and the conditions under which they can justifiable be excluded from the analyzed data set was explored. Finally, we saw that yet another problem that can affect the analysis of random errors is where the measurement set only has a small number of values. In this case, calculations based on the $z$-distribution are inaccurate and we explored the use of a better distribution called the *t-distribution*.

The chapter ended with looking at how the effect of different measurement errors are aggregated together to predict the total error in a measurement system. This process was considered in two parts. First, we looked at how systematic and random error magnitudes can be combined together in an optimal way that best predicts the likely total error in a particular measurement. Second, we looked at situations, where two or more measurements of different quantities are combined together to give a composite measurement value, and looked at the best way of dealing with each of the four arithmetic operations that can be carried out on different measurement components.

## 4.16 Problems

4.1 In a survey of 15 owners of a certain model of car, the following values for average fuel consumption were reported.
25.5 30.3 31.1 29.6 32.4 39.4 28.9 30.0 33.3 31.4 29.5 30.5 31.7 33.0 29.2
Calculate the mean value, the median value, and the standard deviation of the data set.

4.2 The following 10 measurements of the freezing point of aluminum were made using a platinum/rhodium thermocouple.
658.2 659.8 661.7 662.1 659.3 660.5 657.9 662.4 659.6 662.2
Find (a) the median, (b) the mean, (c) the standard deviation, and (d) the variance of the measurements.

4.3 The following 25 measurements were taken of the thickness of steel emerging from a rolling mill.
3.97 3.99 4.04 4.00 3.98 4.03 4.00 3.98 3.99 3.96 4.02 3.99 4.01
3.97 4.02 3.99 3.95 4.03 4.01 4.05 3.98 4.00 4.04 3.98 4.02
Find (a) the median, (b) the mean, (c) the standard deviation, and (d) the variance of the measurements.

4.4 The following 10 measurements are made of the output voltage from a high-gain amplifier that is contaminated due to noise fluctuations:
1.53, 1.57, 1.54, 1.54, 1.50, 1.51, 1.55, 1.54, 1.56, 1.53
Determine the mean value and standard deviation. Hence estimate the accuracy to which the mean value is determined from these 10 measurements. If 1000 measurements were taken, instead of 10, but $\sigma$ remained the same, by how much would the accuracy of the calculated mean value be improved?

4.5 The following measurements were taken with an analog meter of the current flowing in a circuit (The circuit was in steady state and therefore, although the measurements varied due to random errors, the current flowing was actually constant):
21.5, 22.1, 21.3, 21.7, 22.0, 22.2, 21.8, 21.4, 21.9, 22.1 mA
Calculate the mean value, the deviations from the mean, and the standard deviation.

4.6 A batch of digital garden thermometers is tested to determine their accuracy. They are placed in an environment maintained at a temperature of 25 °C, which is measured using an accurate calibrated resistance thermometer. The following readings in °C are obtained from a batch of 10 of the garden thermometers:
24.6 27.2 23.9 26.5 25.4 23.8 26.4 25.9 24.1 25.2
Determine the mean value and standard deviation of these measurements.

4.7 A pressure control system in a distillation column is designed to maintain the pressure at a constant value of 10.9 bar. The following 10 pressure readings (in bars) were obtained at intervals of 10 min:

10.7 11.0 11.2 10.8 10.6 11.1 10.9 10.8 10.7 11.0

Determine the mean and standard deviation of the pressure measurements.

4.8 Using the measurement data given in problem 4.3, draw a histogram of errors (use 5 error bands that are each 0.03 units wide, that is, the center band will be from −0.015 to +0.015).

4.9 (a) Calculate the recommended bin size using the Sturges rule for a histogram of the following set of 25 measurements: 9.4 10.1 9.1 12.3 10.3 10.0 10.5 9.0 10.8 9.9 11.1 9.8 7.6 9.2 10.7 8.4 11.0 9.7 11.3 8.7 9.9 11.5 10.0 9.5 11.9

(b) Draw a histogram of the set of measurements with a bin size of 6:

(c) Draw a histogram of the set of measurements with a bin size of 5:

(d) Which histogram looks most symmetrical?

4.10 (a) When drawing a histogram of measurements, how do you decide on the width of the data bins to use? How do you ensure that there is no ambiguity about which data bin to put any particular measurement into?

(b) Draw a histogram of the following 30 data values which are believed to belong to a Gaussian distribution: 50.0 51.4 49.4 50.6 49.8 50.9 49.2 47.9 49.9 50.7 50.1 49.0 50.6 51.8 49.1 50.6 49.7 49.2 49.8 52.3 50.2 48.3 50.1 50.9 48.9 49.8 49.6 50.3 49.2 50.8

4.11 (a) What do you understand by the term *probability density function*?

Write down an expression for a Gaussian probability density function of given mean value $\mu$ and standard deviation $\sigma$.

4.12 The measurements in a data set are subject to random errors but it is known that the data set fits a Gaussian distribution. Use standard Gaussian tables to determine the percentage of measurements that lie within the boundaries of $\pm 1.5\sigma$, where $\sigma$ is the standard deviation of the measurements.

4.13 The measurements in a data set are subject to random errors but it is known that the data set fits a Gaussian distribution. Use error function tables to determine the value of $x$ required such that 95% of the measurements lie within the boundaries of $\pm x\sigma$, where $\sigma$ is the standard deviation of the measurements.

4.14 By applying error function tables for the mean and standard deviation values calculated in problem 4.3, estimate

(a) How many measurements are <4.00?

(b) How many measurements are <3.95?

(c) How many measurements are between 3.98 and 4.02?

Check your answers against the real data.

4.15 A set of 25 measurements have the following values:
9.4 10.1 9.1 12.3 10.3 10.0 10.5 9.0 10.8 10.0 11.1 9.8 7.6 9.2 10.7 8.4 11.0 9.7 11.3 8.7 9.9 11.5 10.0 9.5 11.9
The mean value and standard deviation of these measurements $x_{mean} = 10.072$ and standard deviation $(\sigma) = 1.1108$.
By applying error function tables to the mean and standard deviation values given, estimate
(a) How many measurements are $<11.05$?
(b) How many measurements are $>9.55$?
(c) How many measurements are between 9.95 and 10.95?
Check your answers against the real data.

4.16 The resolution of the instrument referred to in problem 4.3 is clearly 0.01. Because of the way in which error tables are presented, estimation of the number of measurements in a particular error band is likely to be closer to the real number if the boundaries of the error band are chosen to be between measurement values. In part (c) of problem 3.21, values $>3.98$ are subtracted from values $>4.02$, thus excluding measurements equal to 3.98. Test this hypothesis out by estimating
(a) How many measurements are $<3.995$?
(b) How many measurements are $<3.955$?
(c) How many measurements are between 3.975 and 4.025?
Check your answers against the real data.

4.17 The measurements in a data set are subject to random errors but it is known that the data set fits a Gaussian distribution. Use error function tables to determine the percentage of measurements which lie within the boundaries of $\pm 2\sigma$, where $\sigma$ is the standard deviation of the measurements.

4.18 A silicon integrated circuit chip contains 5000 ostensibly identical transistors. Measurements are made of the current gain of each transistor. The measurements have a mean of 20.0 and a standard deviation of 1.5. The probability distribution of the measurements is Gaussian.
(a) Write down an expression for the number of transistors on the chip, which have a current gain between 19.5 and 20.5.
(b) Show that this number of transistors with a current gain between 19.5 and 20.5 is approximately 1300.
(c) Calculate the number of transistors that have a current gain of 17 or more (this is the minimum current gain necessary for a transistor to be able to drive the succeeding stage of the circuit in which the chip is used).

4.19 In a survey of 12 owners of a certain model of car, the following values for average fuel consumption were reported.
25.5 31.1 29.6 32.4 39.4 28.9 33.3 31.4 29.5 30.5 31.7 29.2

Calculate the mean value, the standard deviation, and the standard error of the mean with the reported consumption values. Express the mean consumption value and the possible error in the mean expressed to 95.4% confidence level.

4.20 In a particular manufacturing process, bricks are produced in batches of 10,000. Because of random variations in the manufacturing process, random errors occur in the target length of the bricks produced. If the bricks have a mean length of 200 mm with a standard deviation of 20 mm, show how the error function tables supplied can be used to calculate the following:
(i) the number of bricks with a length between 198 and 202 mm.
(ii) the number of bricks with a length greater than 170 mm.

4.21 A pressure microsensor is tested by applying a pressure to it of 200 bar, measured by an accurate, calibrated reference pressure-measuring instrument. A set of 12 measurements are made as a reference set of measurements in order to assess the standard deviation and standard error of the mean for measurements made by the device. The measurements obtained for this reference set are given below.
199.7 202.0 200.9 195.7 200.2 199.9 204.4 198.0 203.1 199.1 200.5 196.9
When the microsensor is subsequently used in a workplace to measure the pressure in an enclosed vessel, a reading of 183 bar is obtained. What is the likely error in this measurement, expressed to 95.0% confidence limits?

4.22 The temperature-controlled environment in a hospital intensive care unit is monitored by an intelligent instrument which measures temperature every minute and calculates the mean and standard deviation of the measurements. If the mean is 75 °C and the standard deviation is 2.15,
(a) what percentage of the time is the temperature less than 70 °C?
(b) what percentage of the time is the temperature between 73 °C and 77 °C?

4.23 A semiconductor temperature-measuring device is tested by placing it in an environment that is maintained at a constant temperature of 150 °C, which is measured by an accurate, calibrated resistance thermometer. A set of 12 measurements are made as a reference set of measurements in order to assess the standard deviation and standard error of the mean for measurements made by the device. The measurements obtained for this reference set are given below.
150.7 149.8 150.4 153.5 147.6 148.5 152.4 151.5 149.3 149.9 146.6 150.2
When the sensor is subsequently used in a workplace to measure the temperature in an enclosed vessel, a reading of 137 °C is obtained. What is the likely error in this measurement, expressed to 95.0% confidence limits?

4.24 Calculate the standard error of the mean for the measurements given in problem 4.2. Hence, express the melting point of aluminum together with the possible error in the value expressed to a 68% confidence level.

4.25 The thickness of a set of gaskets varies because of random manufacturing disturbances but the thickness values measured belong to a Gaussian distribution. If the mean thickness is 3 mm and the standard deviation is 0.25, calculate the percentage of gaskets that have a thickness greater than 2.5 mm.

4.26 If the measured variance of 25 samples of breadcakes taken from a large batch is 4.85 g, calculate the true variance of the mass for a whole batch of breadcakes to a 95% significance level.

4.27 Calculate true standard deviation of the diameter of a large batch of tires to a confidence level of 99% if the measured standard deviation in the diameter for a sample of 30 tires is 0.0795 cm.

4.28 150 measurements are taken of the thickness of a coil of rolled steel sheet measured at approximately equidistant points along the center line of its length. The measurements have a mean value of 11.291 mm and a standard deviation of 0.263 mm. The smallest and largest measurements in the sample are 10.73 and 11.89 mm. The measurements are divided into 8 data bins with boundaries at 10.695, 10.845, 10.995, 11.145, 11.295, 11.445, 11.595, 11.745, 11.895. The first bin, containing measurements between 10.695 and 10.845 has 8 measurements in it, and the count of measurements in the following successive bins is 12, 21, 34, 31, 25, 14, 5. Apply the chi-squared test to see whether the measurements fit a Gaussian distribution to a 95% confidence level.

4.29 In a foundry producing castings, the variance in the mass of a sample of 20 castings taken from a large batch is calculated to be 1.36 kg. Calculate the true variance of the mass for a whole batch of castings from which the sample was taken to (a) a 90% significance level and (b) a 95% significance level.

4.30 The temperature in a furnace is regulated by a control system that aims to keep the temperature close to 800 °C. The temperature is measured every minute over a 2-h period, during which time the minimum and maximum temperatures measured are 782 °C and 819 °C. Analysis of the 120 measurements shows a mean value of 800.3 °C and a standard deviation of 7.58 °C. The measurements are divided into 8 data bins of 5 °C width with boundaries at 780.5, 785.5, 790.5, 795.5, 800.5, 805.5, 810.5, 815.5, and 820.5. The measurement count in bin 1 from 780.5 °C to 785.5 °C was 3 and the count in the other successive bins was 8, 21, 30, 28, 19, 9, and 2. Apply the chi-squared test to see whether the measurements fit a Gaussian distribution to (a) a 90% confidence level and (b) a 95% confidence level (NB: think carefully about whether the chi-squared test will be reliable for the measurement counts observed and whether there needs to be any change in the number of data bins used for the chi-squared test).

4.31 Lifetime tests for a sample of 51 tungsten light bulbs taken from a large batch are carried out and the mean lifetime for the sample is found to be 984 h, with a standard deviation of 9.4 h. Calculate the true standard deviation for the whole batch of light bulbs from which the sample of 51 was taken to (a) a 95% significance level and (b) a 99% significance level.

4.32 A pressure control system is designed to maintain the pressure at a constant value of 1000 mbar inside a distillation column.

(a) The following test set of initial pressure measurements is obtained: 1015 1003 1044 1000 985 971 999 990 994 1000 1032 1001 973 1006 979 961 1010 1021 997 1026.
Show that this is free of rogue data points and can therefore be used as a reference set against which to test future pressure measurements.

(b) The pressure is then measured at 5 min intervals and the following 20 measurements are obtained: 961 1010 990 1039 997 1015 1000 949 1000 1051 1001 976 1003 985 932 1006 999 1021 1029 994.
Test whether there are any rogue measurements in this data set.

4.33 (a) What does the *Student variable t* express? Write down an expression for the Student $t$-variable for a small sample ($N$) of measurements in terms of the mean of the sample ($x_{\mathrm{mean}}$), the standard deviation ($\sigma_x$) of the sample, and the mean ($\mu$) of the whole population that the sample is part of in terms of the following ratio:

(b) Sketch typical $t$-distribution curves for sample sizes of 2, 11, and infinity.

(c) Sketch a $t$-distribution curve that shows the meaning of the term $\alpha$, which is commonly used to express the level of confidence that $t$ is less than some value $t_\alpha$.

(d) Use a $t$-distribution table to calculate the value of $t$ for a sample of 21 measurements and a confidence level of 99%.

The measurements are divided into 9 data bins with boundaries at 859.5, 876.5, 893.5, 910.5, 927.5, 944.5, 961.5, 978.5, 995.5, 1012.5.

The measurement count in bin 1 from 869.5 days to 886.5 days is 3 and the count in the other successive bins is 11, 27, 39, 46, 40, 24, 8, and 2.

4.34 Accelerated lifetime tests are carried out for a batch of 200 thermocouples. The mean lifetime is found to be 936 h with a standard deviation of 28 h. The smallest and largest measurements in the sample are 860 and 1012 h. The measurements are divided into 9 data bins with boundaries at 859.5, 876.5, 893.5, 910.5, 927.5, 944.5, 961.5, 978.5, 995.5, 1012.5. The measurement count in bin 1 from 869.5 days to 886.5 days is 3 and the count in the other successive bins is 8, 27, 39, 46, 40, 24, 11, and 2. Apply the chi-squared test to see whether the measurements fit a Gaussian distribution to (a) a 99% confidence level, (b) a 97.5% confidence level, and (c) a 95% confidence level.

4.35 A temperature control system is designed to maintain the temperature at a constant value of 500 °C inside an induction furnace.

(a) The following test set of initial temperature measurements (once the furnace temperature has reached steady state) is obtained: 507 502 520 500 492 485 499 495 497 500 516 501 489 503 490 481 504 511 499 512

Show that this is free of rogue data points and can therefore be used as a reference set against which to test future temperature measurements.

(b) The temperature is then measured at 10 min intervals and the following 20 measurements are obtained:

481 505 495 519 498 1015 508 477 500 526 501 488 502 492 464 503 499 510 515 497

Test whether there are any rogue measurements in this data set.

4.36 The internal diameter of a large batch of pressure vessels is measured by destructive testing of 8 samples taken randomly from the batch. If the sample mean is 346 mm with a standard deviation of 7 mm, express the upper and lower bounds to a confidence level of 95% on the range in which the mean value lies for internal diameter of the whole batch.

4.37 The volume contained in each of a sample of 10 bottles of expensive perfume is measured. If the mean volume of the sample measurements is 100.5 ml with a standard deviation of 0.64 ml, calculate the upper and lower bounds to a confidence level of 95% of the mean value of the whole batch of perfume from which the 10 samples were taken.

4.38 A 3-V DC power source required for a circuit is obtained by connecting together two 1.5 V batteries in series. If the error in the voltage output of each battery is specified as $\pm 1\%$, calculate the likely maximum error in the 3-V power source that they make up.

4.39 A temperature measurement system consists of a thermocouple, whose amplified output is measured by a voltmeter. The output relationship for the thermocouple is approximately linear over the temperature range of interest. The emf/temp relationship of the thermocouple has a possible error of $\pm 1\%$, the amplifier gain value has a possible error of $\pm 0.5\%$, and the voltmeter has a possible error of $\pm 2\%$. What is the possible error in the measured temperature?

4.40 A pressure measurement system consists of a monolithic piezoresistive pressure transducer and a bridge circuit to measure the resistance change of the transducer. The resistance ($R$) of the transducer is related to pressure ($P$) according to $R = K_1 P$ and the output of the bridge circuit ($V$) is related to resistance ($R$) by $V = K_2 R$. Thus, the output voltage is related to pressure according to $V = K_1 K_2 P$.

If the maximum error in $K_1$ is $\pm 2\%$, the maximum error in $K_2$ is $\pm 1.5\%$, and the voltmeter itself has a maximum measurement error of is $\pm 1\%$, what is the likely maximum error in the pressure measurement?

4.41 A requirement for a resistance of 1220 Ω in a circuit is satisfied by connecting together resistances of 1000 and 220 Ω in series. If each resistance has a tolerance of $\pm 5\%$, what is the likely tolerance in the total resistance?

4.42 In order to calculate the heat loss through the wall of a building, it is necessary to know the temperature difference between the inside and outside walls. Temperatures of 5 °C and 20 °C are measured on each side of the wall by mercury-in-glass thermometers with a range of 0 °C to +50 °C and a quoted inaccuracy of $\pm 1\%$ of full-scale reading.

Calculate the likely maximum possible error in the calculated value for the temperature difference.

Briefly discuss how using measuring instruments with a different measurement range might improve measurement accuracy.

4.43 A fluid flow rate is calculated from the difference in pressure measured across a venturi. Flow rate is given by $F = K(p_2 - p_1)$.

where $p_1$, $p_2$ are the pressures at either side of the venturi and $K$ is a constant.

The two pressure measurements are 15.2 bar and 14.7 bar.

(a) Calculate the possible error in flow measurement if the pressure-measuring instruments have a quoted error of $\pm 0.2\%$ of their reading.

(b) Briefly discuss why using a differential pressure sensor rather than two separate pressure sensors would improve measurement accuracy.

4.44 The power dissipated in a car headlight is calculated by measuring the DC voltage drop across it and the current flowing through it ($P = V \times I$). If the possible errors in the measured voltage and current values are $\pm 1\%$ and $\pm 2\%$ respectively, calculate the likely maximum possible error in the power value deduced.

4.45 The resistance of a carbon resistor is measured by applying a DC voltage across it and measuring the current flowing ($R = V/I$). If the voltage and current values are measured as $10 \pm 0.1$ V and $214 \pm 5$ mA respectively, express the value of the carbon resistor.

4.46 The specific energy of a substance is calculated by measuring the energy content of a cubic meter volume of the substance. If the errors in energy measurement and volume measurement are $\pm 1\%$ and $\pm 2\%$ respectively, what is the possible error in the calculated value of specific energy? (NB: specific energy = energy per unit volume of material).

4.47 (a) In a particular measurement system, a quantity $x$ is calculated by subtracting a measurement of a quantity $z$ from a measurement of a quantity $y$, that is, $x = y - z$. If the possible measurement errors in $y$ and $z$ are $\pm ay$ and $\pm bz$ respectively, show that the value of $x$ can be expressed as

$$x = y - z \pm (ay - bz)$$

What is inconvenient about this expression for $x$ and what is the basis for the following expression for $x$ which is used more commonly?

$$x = (y - z) \pm e$$

where $e = \sqrt{(ay)^2 + (bz)^2}$.

(b) In a different measurement system, a quantity $p$ is calculated by multiplying together measurements of two quantities $q$ and $r$ such that $p = qr$. If the possible measurement errors in $q$ and $r$ are $\pm aq$ and $\pm br$, respectively, show that the value of $p$ can be expressed as

$$p = (qr)(1 \pm [a + b])$$

(c) The volume flow rate of a liquid in a pipe (the volume flowing in unit time) is measured by allowing the end of the pipe to discharge into a vertical-sided tank with a rectangular base (see Figure 4.13). The depth of the liquid in the tank is measured at the start as $h_1$ meters and 1 min later it is measured as $h_2$ meters. If the length and width of the tank are $l$ meters and $w$ meters respectively, write down an expression for the volume flow rate of the liquid in cubic meters per minute.
Calculate the volume flow rate of the liquid if the measured parameters have the following values:

$$h_1 = 0.8 \text{ m}; \quad h_2 = 1.3 \text{ m}; \quad l = 4.2 \text{ m}; \quad w = 2.9 \text{ m}$$

If the possible errors in the measurements of $h_1$, $h_2$, $l$, and $w$ are 1%, 1%, 0.5%, and 0.5% respectively, calculate the possible error in the calculated value of the flow rate.

4.48 The density of a material is calculated by measuring the mass of a rectangular-sided block of the material, whose edges have lengths of $a$, $b$, $c$. What is the possible error in the calculated density if the possible error in mass measurement is $\pm 1.0\%$ and the possible errors in length measurement are $\pm 0.5\%$?

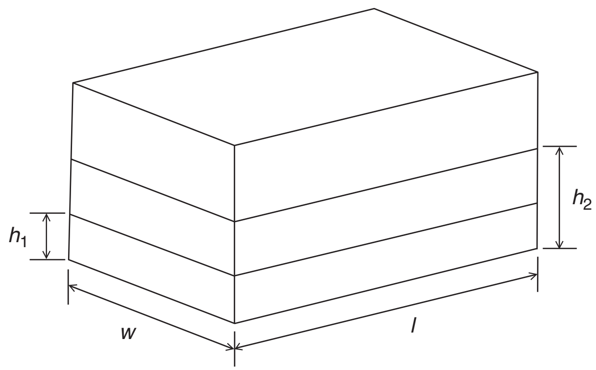


**Figure 4.13**
Diagram for problem 4.47.

4.49 The density ($d$) of a liquid is calculated by measuring its depth ($c$) in a calibrated rectangular tank and then emptying it into a mass measuring system. The length and width of the tank are ($a$) and ($b$) respectively and thus the density is given by

$$d = m/(a \times b \times c)$$

where $m$ is the measured mass of the liquid emptied out.
If the possible errors in the measurements of $a$, $b$, $c$, and $m$ are 1%, 1%, 2%, and 0.5%, respectively, determine the likely maximum possible error in the calculated value of the density ($d$).

4.50 The volume flow rate of a liquid is calculated by allowing the liquid to flow into a cylindrical tank (stood on its flat end) and measuring the height of the liquid surface before and after the liquid has flowed for 10 min. The volume collected after 10 min is given by

$$\text{Volume} = (h_2 - h_1)\pi(d/2)^2$$

where $h_1$ and $h_2$ are the starting and finishing surface heights and $d$ is the measured diameter of the tank.
(a) If $h_1 = 2$ m, $h_2 = 3$ m, and $d = 2$ m, calculate the volume flow rate in m$^3$/min.
(b) If the possible error in each measurement $h_1$, $h_2$, and $d$ is $\pm 1\%$, determine the likely maximum possible error in the calculated value of volume flow rate. (It is assumed that there is negligible error in the time measurement.)

## References

ANSI/ASME standards, 1998. PTC19.1-ASME Performance Test Codes: Measurement Uncertainty. American Society of Mechanical Engineers, New York.

Topping, J., 1972. Errors of Observation and Their Treatment. Chapman and Hall (N.B. this is a long-established authoritative guide to the subject that is still in print).