10 Probabilities and Errors

If there are different outcomes to the performance of a procedure, there is a set of probabilities for the different outcomes. There are six possible results if a die is thrown. If the die is unbiased, the probability of any integer from one to six being face-up is the same for any of the six numbers. This is not the usual situation. The probabilities for different horses winning a race are normally different. Whether the possibilities of a procedure are finite or to all intents and purposes limitless, such as when the outcomes form a continuous set, the sum of the probabilities for all outcomes is unity. There must be a result to the procedure. A working definition of the probability of a certain result from any experiment or procedure is to imagine doing the identical experiment many times and noting the fraction of times the particular result is obtained. This is the probability, $P(n)$, of the occurrence of the result n.

For many procedures, or many similar types of procedure where the result can be expressed as a number n, the distribution of probabilities is known from experience, and can be expressed as a function of n. This function is called the **distribution function**. All experimental measurements have random errors which arise from many causes, such as the random imprecision of the measuring devices and the imperfections of the person making the measurements. One measurement of a quantity which is believed to have a 'true', constant value, such as the length of a rod, is different from the previous measurement and the deviation of any measurement from the 'true' value may be on either side of the 'true' value independently of the sign of the previous deviation. The observed lengths of the rod are distributed about a most probable value and the manner in which they are distributed is described by the distribution function appropriate to random errors of measurement. If the errors are small, and many measurements of the length of the rod are made, the spread of the observations is smaller than if the errors are large.

10.1 Distribution functions

Some quantities relevant to distribution functions are outlined in this sub-section before discussing two of the most useful distributions used in physics. Let us retain the example of the length of a rod. If the length is measured several times, and the number of times, $N(L)$ a length between $L$and $L + \Delta L$ is recorded is plotted against L, the result looks like Figure 10.1. The measurements have been put in bins of width ΔL to produce a **histogram** of the data.

For a reasonably large number of measurements, the histogram is roughly symmetrical about the bin containing the most probable length measured. If very many measurements are made, the increased number of data points will allow narrower bins to be used in the histogram. Eventually the distribution becomes a continuous curve. The function $P(L)$ describing the continuous curve is the distribution function of the measured lengths and the probability that the length lies between L and $L + \mathrm{d}L$ is $P(L)\mathrm{d}L$. Note that in this example $P(L)$ has the dimensions of a reciprocal length.

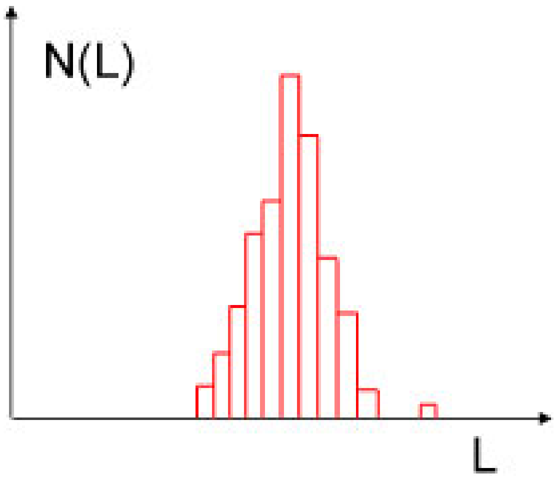

Figure 10.1

The **moments** of the distribution are properties of the distribution which have important uses. The first moment is the **mean** or the **average**. If N measurements of the length are taken and the i'th measurement gives the value L_i, the average $< L >$ is

$$< L >= \frac{\sum_i L_i}{N}. \tag{10.1}$$

The **variance** of the distribution of the lengths is a measure of the spread about the mean, and is defined as

$$V = \frac{\sum_i (L_i - < L >)^2}{N}. \tag{10.2}$$

An equivalent formula for the variance is obtained from the above by expanding

$$\begin{aligned} V &= \frac{1}{N}\sum_i L_i^2 - \frac{1}{N}\sum_i 2L_i < L > + \frac{1}{N}\sum_i < L >^2 \\ &= \frac{1}{N}\sum_i L_i^2 - 2 < L > \frac{1}{N}\sum_i L_i + < L >^2 \\ &= < L^2 > -2(< L >)^2 + (< L >)^2 = < L^2 > - < L >^2 . \end{aligned} \tag{10.3}$$

The **standard deviation** is most often used as a measure of the spread of a distribution. It is simply the square root of the variance and has the symbol σ.

$$\sigma = \sqrt{V}. \tag{10.4}$$

Example 10.1 Fifteen measurements are made of the length of a rod. The values obtained in cm are 24.1, 24.2, 24.4, 24.3, 24.3, 24.2, 24.5, 24.4, 24.3, 24.2, 24.5, 24.4, 24.3, 24.4, and 24.2. What is the mean, the variance and standard deviation of the set of results?

Solution From equation 10.1, the mean is 364.7/15=24.3133 cm. The average of the square of the lengths is 8867.27/15=591.151 cm^2. From equation 10.3, the variance is then 591.151-591.138=0.013 cm^2. The standard deviation is the square root of the variance and equals 0.114 cm.

Problem 10.1 In the above example, what is the percentage change to the mean and the standard deviation of the addition of an unlikely measurement of 25.3 cm?

For a continuous distribution $P(x)$ of a quantity x, the mean or average value of x is

$$< x >= \frac{\int_{-\infty}^{\infty} xP(x)\mathrm{d}x}{\int_{-\infty}^{\infty} P(x)\mathrm{d}x}. \tag{10.5}$$

In the above formula we have written the lower limit of integration to be $-\infty$ to cover the general situation in which negative values are possible.

For distributions which are symmetrical about zero, the mean is zero. The first non-vanishing moment of the distribution is then the **mean square** value. The mean square of x is

$$< x^2 >= \frac{\int_{-\infty}^{\infty} x^2P(x)\mathrm{d}x}{\int_{-\infty}^{\infty} P(x)\mathrm{d}x}. \tag{10.6}$$

Higher moments of the distribution may be defined in a similar fashion. For a continuous distribution function $P(x)$ the variance is

$$V = \frac{\int_{-\infty}^{\infty}(x- < x >)^2P(x)\mathrm{d}x}{\int_{-\infty}^{\infty} P(x)\mathrm{d}x}. \tag{10.7}$$

With the total probability normalised to unity the denominators in the above expressions 10.5 to 10.7 reduce to unity.

10.2 The Gaussian distribution We now consider the distribution function which random errors of measurements obey. The deviations of measurements of a quantity from its true value are as likely to be positive as negative, and if many measurements are taken, the distribution is symmetrical about the most probable measured value. Their distribution function is a **Gaussian** curve. A Gaussian curve is shown in Figure 10.2. The mathematical expression for the curve when it has a maximum at x_0 is

$$P(x) = P_0 \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right), \tag{10.8}$$

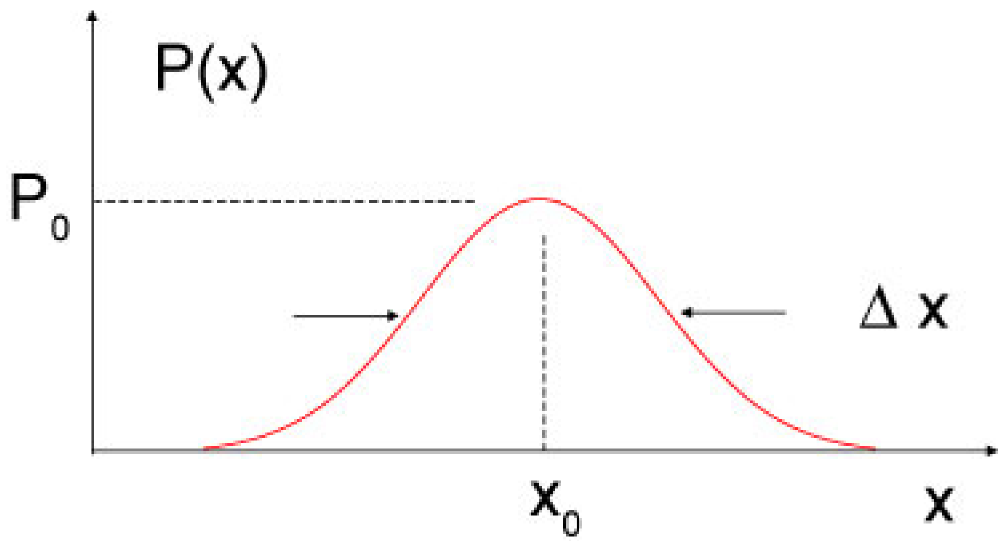

Figure 10.2

where P_0 is the maximum value of P and the parameter σ is related to the full width Δx of the curve at one half its maximum height P_0.

The variance of a Gaussian distribution function can be evaluated using the definition (10.7) and equals the square of the parameter σ.

$$V = \sigma^2. \tag{10.9}$$

Hence the standard deviation of a Gaussian distribution is equal to the parameter σ.

Example 10.2 show that $\Delta x = 2\sigma\sqrt{2\ln 2} = 2.355\sigma$, to four significant figures. *Solution* For the Gaussian function, when $P(x) = P_0/2$,

$$\exp\left(+\frac{(x-x_0)^2}{2\sigma^2}\right) = 2$$

and $(x-x_0)^2 = 2\sigma^2 \ln 2$. Hence $x = x_0 \pm \sqrt{2\sigma^2 \ln 2}$ and

$$\Delta x = 2\sigma\sqrt{2\ln 2} = 2.355\sigma.$$

Problem 10.2 Use values of definite integrals given in Appendix A to prove equation 10.9.

The integral from minus infinity to plus infinity of a Gaussian function with peak height unity is

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right)\, \mathrm{d}x = \sqrt{2\pi}\ \sigma. \tag{10.10}$$

The total probability is made equal to unity by putting P_0 in equation 10.8 equal to $1/\sqrt{2\pi}\ \sigma$ as dictated by the above equation.

10.3 The Poisson distribution

The Poisson distribution describes the frequency with which random, rare events occur. The number of decays from a radioactive source follows a Poisson distribution. If equation 7.2 is used to determine the number of decays in a given time interval it predicts the average number. The actual number will fluctuate around the mean, with the percentage spread smaller the larger the mean value. A small amount of radioactive material of exceedingly long half-life may undergo only a few decays a month. The decays are rare events and their likelihood follows a Poisson distribution. If the average number of decays per month over a long period is $< n >$, what is the probability of n decays in a particular month's observation?

This probability $P(n)$ is given by

$$P(n, < n >) = \frac{\mathrm{e}^{-<n>}\ (< n >)^n}{n!}. \tag{10.11}$$

where $n!$ equals $n(n-1)(n-2)....1$ and is factorial n. Note that n is an integer but $< n >$ will normally be non-integral.

Example 10.3 If the average number of decays of a radioactive source per day is 1.2, what is the probability of the occurrence of 4 such events in a particular day's study? *Solution* The average $< n >= 1.2$. The probability $P(n, < n >)$ for $n = 4$ is

$$P(4, 1) = \frac{e^{-1.2} 1.2^4}{4!} = 0.026.$$

Problem 10.3 Show that the total probability of all possible numbers of events occurring, $\sum_{n=0}^{n=\infty} P(n)$, is unity.

The variance of a Poisson distribution is $< n >$ and the standard deviation of the mean is $\sqrt{< n >}$. This can be proved by noting that

$$< n(n-1) >= \sum_{n=0}^{\infty} n(n-1) e^{-<n>} \frac{< n >^n}{n!},$$

where we have used the formula for the average of a function $f(n)$ in terms of the probability distribution, $P(n)$, for n. The first two terms on the right-hand side of the above equation are zero and the right-hand side becomes

$$< n(n-1) > = < n >^2 \sum_{n=2}^{\infty} e^{-<n>} \frac{< n >^{n-2}}{(n-2)!}$$

$$= < n >^2 \sum_{m=0}^{\infty} e^{-<n>} \frac{< n >^{m}}{m!},$$

putting $(n-2) = m$.

The sum in the last equation is unity, from Example 10.2. Hence

$$< n(n-1) > = (< n >)^2 = < n^2 > - < n >,$$

and

$$< n^2 > = (< n >)^2 + (< n >).$$

Thus the variance, given by equation 10.2, is

$$V = < n^2 > - (< n >)^2$$

$$= (< n >)^2 + (< n >) - (< n >)^2 = (< n >), \tag{10.12}$$

and the standard deviation is

$$\sigma = \sqrt{V} = \sqrt{< n >}. \tag{10.13}$$

If the mean number of decays observed over a given time interval, $< n >$, is large, as it may be for radioactive decays from sources usually used in the laboratory, the Poisson distribution becomes closely equal to the Gaussian distribution. If the number of counts observed in a particular time interval is N, N is close to the mean of a sample of counts, and the error on N is $\sqrt{N}$ and the fractional error on the count is $1/\sqrt{N}$.

Problem 10.4 Neutrino bursts from a galaxy contain a mean of 12 neutrinos per burst. Hypothetical detection equipment first converts neutrinos to electrons with an efficiency of 0.1% and is then able to record single electrons or more. What is the probability that the equipment will detect a burst?

10.4 Errors

All experimental measurements have errors, both random and systematic.

The random errors follow a Gaussian distribution given by equation 10.8. The measurements of a quantity x are distributed about the most probable value x_0 and the probability of the occurrence of a measured value between x and $x + \mathrm{d}x$ is equal to $P(x)\mathrm{d}x$ when P is normalised to make the total probability unity. This is done by putting P_0 in equation 10.8 equal to $1/\sqrt{2\pi}\ \sigma$ as dictated by equation 10.10. However, the deviations from the 'true' value occur randomly, and for a finite number of measurements in a set or sample the mean differs from the mean of a similar set taken before or after.

If a sample contains an infinite number of measurements, their distribution is a continuous curve and the mean is known accurately. However, an infinite number of measurements can not be made and the determination of the 'true' value from the mean of a limited set has an uncertainty which depends on the number of data points taken. The mean of the measurements is usually quoted as the best value determined, and the error is given as the **standard error on the mean** which is

$$\text{standard error on the mean} = \sigma/\sqrt{N}, \tag{10.14}$$

where σ is the standard deviation of the Gaussian distribution function of the set of points. It is important to recognise the distinction between the standard deviation of the distribution function which represents the spread of measured values, from the accuracy with which the mean is known.

- **Propagation of errors**

A measured quantity z may depend upon two other measured quantities x and y through a functional relationship

$$z = f(x, y).$$

A small change Δz in z is produced by small changes Δx and Δy in x and y, with

$$\Delta z = \left(\frac{\partial f}{\partial x}\right)\ \Delta x + \left(\frac{\partial f}{\partial y}\right)\ \Delta y.$$

Squaring both sides gives

$$(\Delta z)^2 = \left(\frac{\partial f}{\partial x}\right)^2\ (\Delta x)^2 + \left(\frac{\partial f}{\partial y}\right)^2\ (\Delta y)^2 + 2\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right)\ \Delta x\ \Delta y.$$

If the small increment Δz is the small error introduced by errors Δx and Δy, and if the latter errors are uncorrelated and equally likely to be negative as positive, over many measurements the cross-term involving $\Delta x \Delta y$ in the above equation averages to zero. Regarding Δx and Δy as standard deviations on the measurements of x and y, and Δz as the consequent standard deviation in the measured value of z, we then have

$$\sigma_z^2 = \left(\frac{\partial f}{\partial x}\right)^2\ \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2\ \sigma_y^2. \tag{10.15}$$

Example 10.4 Ohm's Law states that $V = IR$. What is the error on the voltage V when there is a measured current of 0.7 ± 0.02 A through a resistance of 10 ± 0.2 Ω?

Solution

$$\sigma_V^2 = \left(\frac{\partial V}{\partial I}\right)^2 \sigma_I^2 + \left(\frac{\partial V}{\partial R}\right)^2 \sigma_R^2$$

$$= R^2\sigma_I^2 + I^2\sigma_R^2$$

$$= 100.(0.02)^2 + 0.7^2(0.2)^2 = 0.0596.$$

hence $\sigma_V = 0.244$ V.

Problem 10.5 What is the current through the resistor of the above example when the voltage across it is measured to be 12 ± 0.2 V?

- **Systematic errors**

There are often systematic errors present in a measurement. For example, a ruler used to measure the length of a rod may have a calibration error. It may consistently read too high or too low by a given amount but we don't know by how much and in what direction. In that case, the systematic error has to be combined with the random error on the mean, obtained from the spread in the length readings using equations 10.9 and 10.11, to give a total error.

Let the ruler's calibration be uncertain to $\pm\sigma_S$. If the mean length measured is $< L >$ and the standard error on the mean is σ_M, the mean value of the length has the additional systematic error σ_S. The systematic error is usually combined with the random error by adding them in **quadrature**. The total standard deviation σ is given by

$$\sigma^2 = \sigma_M^2 + \sigma_S^2, \tag{10.16}$$

and the length is quoted as

$$L =< L > \pm\sigma.$$

Sometimes the random and systematic errors are quoted separately and the measurement given as

$$L =< L > \pm\sigma_M \pm \sigma_S. \tag{10.17}$$

10.5 Least-squares fitting

Many experiments measure the value of one quantity y as a second x is varied. For example, y may be the length of a metal rod and x may be the temperature at which the length is measured. If we assume there are only small errors on the measurements of temperature and that these errors can be neglected compared with the errors on the measured lengths, the data may be as shown on Figure 10.3.

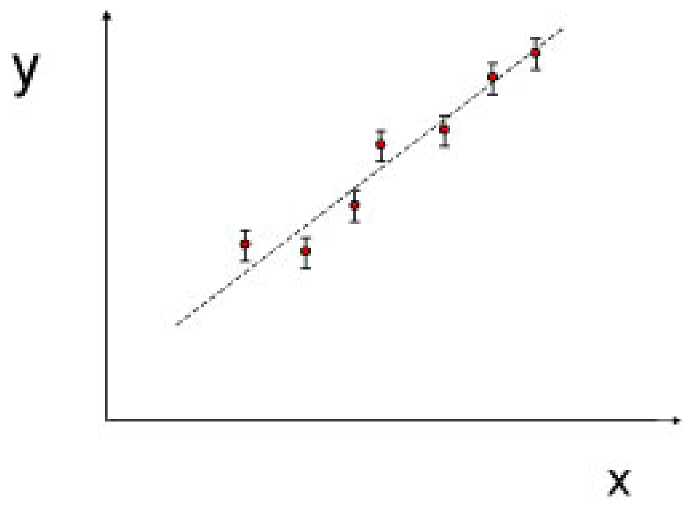

Figure 10.3

Over a limited temperature range there is reason to believe that y varies linearly with x over the temperature range considered. Hence, over that range, y is given by the formula

$$y = a + bx. \tag{10.18}$$

Because of the errors on the points, a line given by equation (10.18) is unlikely to pass through all the points on Figure 10.3 whatever line is chosen. The problem is to determine the best line from the data set. The best values of the parameters a and b are given by **fitting** the points to a straight line of the form given by equation (10.18).

Let a measurement y_i of y have an error Δy_i. We define a quantity called the **chi-squared**, symbol χ^2, given by

$$\chi^2(a,b) = \sum_{i=1,N} \frac{(y_i - y(a,b))^2}{\Delta y_i^2}, \tag{10.19}$$

where the sum is over the N pairs of (x, y), and $y(a, b)$ is the value predicted by equation (10.18) using parameter values a and b. The best values of a and b are then those which correspond to the minimum value of χ^2.

This technique of estimating the best parameter values is called **least-squares fitting**. The best values when all y_i have the same errors may be shown to be given by

$$a = \frac{< x^2 > \ < y > - < x > \ < xy >}{< x^2 > - < x >^2}, \tag{10.20}$$

and

$$b = \frac{< xy > - < x > \ < y >}{< x^2 > - < x >^2}. \tag{10.21}$$

The standard deviations on these parameters are

$$\sigma_b^2 = \frac{\sigma_y^2}{N(< x^2 > - < x >^2)}, \tag{10.22}$$

from equation 10.15, and similarly,

$$\sigma_a^2 = \frac{\sigma_y^2 < x^2 >}{N(< x^2 > - < x >^2)}. \tag{10.23}$$

Problem 10.6 Seven values of the resistance R of a wire are made at different temperatures T. The errors on the temperature measurements are negligible and the errors on the resistance measurements are all $\pm 0.1\Omega$. Corresponding data points (temperature first in degrees Kelvin and resistance in ohms) are: (280, 6.9), (300, 7.1), (320, 7.0), (340, 7.5), (360, 7.5), (380, 7.9) and (400, 7.9). Show that the slope of the best-fit straight line through the points is 0.0102 (ΩK^{-1}).

Problem 10.7 Using the above values for a and b show that the best fit straight line through a set of points (x, y) passes through the point $(< x >, < y >)$.
