Hence, using eqn 3.22, the variance in y is given by eqn 3.27 minus eqn 3.28, i.e.

$$\begin{aligned}\sigma_y^2 &= \langle y^2\rangle - \langle y\rangle^2 \\ &= a^2\langle x^2\rangle - a^2\langle x\rangle^2 \\ &= a^2\sigma_x^2. \end{aligned} \quad (3.29)$$

Notice that the variance depends on a but not on b. This makes sense because the variance tells us about the width of a distribution, and nothing about its absolute position. The standard deviation of y is therefore given by

$$\sigma_y = a\sigma_x. \quad (3.30)$$

Example 3.6

The average temperature in a town in the USA in January is 23 °F and the standard deviation is 9 °F. Convert these figures into degrees Celsius using the relation in Example 3.4.

Solution:

The average temperature in degrees Celsius is given by

$$\langle C\rangle = \frac{5}{9}(\langle F\rangle - 32) = \frac{5}{9}(23-32) = -5\,°\mathrm{C}, \quad (3.31)$$

and the standard deviation is given by $\frac{5}{9}\times 9 = 5\,°\mathrm{C}$.

3.6 Independent variables

If u and v are **independent random variables**[4] the probability that u is in the range from u to $u+\mathrm{d}u$ and v is in the range from v to $v+\mathrm{d}v$ is given by the product

$$P_u(u)\mathrm{d}u\, P_v(v)\mathrm{d}v. \quad (3.32)$$

Hence, the average value of the product of u and v is

$$\begin{aligned}\langle uv\rangle &= \iint uvP_u(u)P_v(v)\,\mathrm{d}u\,\mathrm{d}v \\ &= \int uP_u(u)\,\mathrm{d}u\int vP_v(v)\,\mathrm{d}v \\ &= \langle u\rangle\langle v\rangle, \end{aligned} \quad (3.33)$$

because the integrals separate for *independent* random variables. This implies that the average value of the product of u and v is equal to the product of their average values.

[4] Two random variables are independent if knowing the value of one of them yields no information about the value of the other. For example, the height of a person chosen at random from a city and the number of hours of rainfall in that city on the first Tuesday of September are two independent random variables.

Example 3.7

Suppose that there are n independent random variables, X_i, each with the same mean $\langle X\rangle$ and variance σ_X^2. Let Y be the sum of the random variables, so that $Y = X_1 + X_2 + \cdots + X_n$. Find the mean and variance of Y.

Solution:

The mean of Y is simply

$$\langle Y\rangle = \langle X_1\rangle + \langle X_2\rangle + \cdots + \langle X_n\rangle, \quad (3.34)$$

but since all the X_i have the same mean $\langle X\rangle$ this can be written

$$\langle Y\rangle = n\langle X\rangle. \quad (3.35)$$

Hence the mean of Y is n times the mean of the X_i. To find the variance of Y, we can use the formula

$$\sigma_Y^2 = \langle Y^2\rangle - \langle Y\rangle^2. \quad (3.36)$$

Hence

$$\begin{aligned}\langle Y^2\rangle &= \langle X_1^2 + \cdots + X_N^2 + X_1X_2 + X_2X_1 + X_1X_3 + \cdots\rangle \quad (3.37)\\ &= \langle X_1^2\rangle + \cdots + \langle X_N^2\rangle + \langle X_1X_2\rangle + \langle X_2X_1\rangle + \langle X_1X_3\rangle + \cdots \end{aligned}$$

There are n terms like $\langle X_1^2\rangle$ on the right-hand side, and $n(n-1)$ terms like $\langle X_1X_2\rangle$. The former terms take the value $\langle X^2\rangle$ and the latter terms (because they are the product of two independent random variables) take the value $\langle X\rangle\langle X\rangle = \langle X\rangle^2$. Hence, using eqn 3.35,

$$\langle Y^2\rangle = n\langle X^2\rangle + n(n-1)\langle X\rangle^2, \quad (3.38)$$

so that

$$\begin{aligned}\sigma_Y^2 &= \langle Y^2\rangle - \langle Y\rangle^2 \\ &= n\langle X^2\rangle - n\langle X\rangle^2 \\ &= n\sigma_X^2. \end{aligned} \quad (3.39)$$

The results proved in this last example have some interesting applications. The first concerns experimental measurements. Imagine that a quantity X is measured n times, each time with an independent error, which we call σ_X. If you add up the results of the measurements to make $Y = \sum X_i$, then the rms error in Y is only $\sqrt{n}$ times the rms error of a single X. Hence if you try and get a good estimate of X by calculating $(\sum X_i)/n$, the error in this quantity is equal to $\sigma_X/\sqrt{n}$. Thus, for example, if you make four measurements of a quantity and average your results, the random error in your average is half of what it would be if you'd just taken a single measurement. Of course, you may still have *systematic* errors in your experiment. If you are consistently overestimating your quantity by an error in your experimental setup, that error won't reduce by repeated measurement!

A second application is in the theory of **random walks**. Imagine a drunken person staggering out of a pub and attempting to walk along a narrow street (which confines him or her to motion in one dimension). Let's pretend that with each inebriated step, the drunken person is equally likely to travel one step forwards or one step backwards. The effects of intoxication are such that each step is uncorrelated with the previous one. Thus the average distance travelled in a single step is $\langle X \rangle = 0$. After n such steps, we would have an expected total distance travelled of $\langle Y \rangle = \sum \langle X_i \rangle = 0$. However, in this case the root mean squared distance is more revealing. In this case $\langle Y^2 \rangle = n\langle X^2 \rangle$, so that the rms length of a random walk of n steps is $\sqrt{n}$ times the length of a single step. This result will be useful in considering Brownian motion in Chapter 33.

3.7 Binomial distribution

A probability distribution, which is very important in thermal physics, is based on what is called a **Bernoulli trial**,[5] an "experiment" with two possible outcomes. One outcome (which we will call "success") occurs with probability p and the other outcome (which we will call "failure") occurs with probability $1 - p$. An example of a Bernoulli trial is the tossing of a coin: one outcome is "heads", the other is "tails".

[5] Jacob Bernoulli (1654–1705).

Example 3.8

Let x be a random variable which takes the value 1 for success and 0 for failure. Then, assuming p to be the probability of success and using eqns 3.2, 3.3 and 3.21

$$\begin{aligned} \langle x \rangle &= 0 \times (1-p) + 1 \times p = p && (3.40) \\ \langle x^2 \rangle &= 0^2 \times (1-p) + 1^2 \times p = p && (3.41) \\ \sigma_x &= \sqrt{\langle x^2 \rangle - \langle x \rangle^2} = \sqrt{p(1-p)}. && (3.42) \end{aligned}$$

The **binomial distribution** is the discrete probability distribution $P(n, k)$ of getting k successes from n independent Bernoulli trials. The function $P(n, k)$ can be worked out by realizing that (a) the probability of a particular series of k successes and $n-k$ failures is $p^k(1-p)^{n-k}$ and (b) that there are ^nC_k ways of arranging k successes and $n-k$ failures

The **binomial theorem** of elementary algebra states that

$$(x+y)^n = \sum_{k=0}^{n} {^nC_k}\, x^k y^{n-k}. \qquad (3.44)$$

Hence by writing $x = p$ and $y = 1 - p$ we can easily show that

$$\sum_{k=1}^{n} P(n,k) = 1, \qquad (3.45)$$

as required for a well-behaved probability distribution. Since the binomial distribution is the sum of n *independent* Bernoulli trials, then

$$\begin{aligned} \langle k \rangle &= np && (3.46) \\ \sigma_k^2 &= np(1-p). && (3.47) \end{aligned}$$

The **fractional width** of the distribution[6] is obtained by dividing the standard deviation by the mean and is given by $\sigma_k / \langle k \rangle = \sqrt{(1-p)/np}$, which is proportional to $1/\sqrt{n}$, and therefore decreases as n increases. This causes the binomial distribution to become more sharply peaked near the mean value as n increases, as shown in Fig. 3.3.

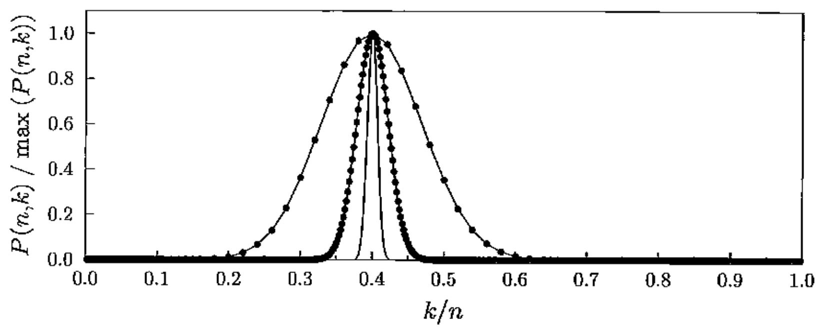

Fig. 3.3 Binomial probability for $p = 0.4$. The three plots are for $n = 50$ (outermost), $n = 500$ and $n = 5000$ (innermost) and are scaled so that their maximum amplitudes are the same. This demonstrates that as n increases, the *fractional width* decreases.

Example 3.9

Coin tossing with a fair coin. In this case, $p = \frac{1}{2}$.

- For $n = 16$ tosses, the expected number of heads is $np = 8$. The standard deviation is $\sqrt{np(1-p)} = 2$, a quarter of the expected number.
- For $n = 10^{20}$ tosses, the expected number of heads is $np = 5 \times 10^{19}$.