where the average is weighted by the relative probabilities of the different values of x. For continuous variables, expectations are expressed in terms of an integration with respect to the corresponding probability density:

$$\mathbb{E}[f] = \int p(x) f(x) \, \mathrm{d}x. \tag{2.39}$$

In either case, if we are given a finite number N of points drawn from the probability distribution or probability density, then the expectation can be approximated as a finite sum over these points:

Exercise 2.7

$$\mathbb{E}[f] \simeq \frac{1}{N} \sum_{n=1}^{N} f(x_n). \tag{2.40}$$

The approximation in (2.40) becomes exact in the limit $N \to \infty$.

Sometimes we will be considering expectations of functions of several variables, in which case we can use a subscript to indicate which variable is being averaged over, so that for instance

$$\mathbb{E}_x[f(x, y)] \tag{2.41}$$

denotes the average of the function $f(x, y)$ with respect to the distribution of x. Note that $\mathbb{E}_x[f(x, y)]$ will be a function of y.

We can also consider a *conditional expectation* with respect to a conditional distribution, so that

$$\mathbb{E}_x[f|y] = \sum_x p(x|y) f(x), \tag{2.42}$$

which is also a function of y. For continuous variables, the conditional expectation takes the form

$$\mathbb{E}_x[f|y] = \int p(x|y) f(x) \, \mathrm{d}x. \tag{2.43}$$

The *variance* of $f(x)$ is defined by

$$\mathrm{var}[f] = \mathbb{E}\left[\left(f(x) - \mathbb{E}[f(x)] \right)^2 \right] \tag{2.44}$$

and provides a measure of how much $f(x)$ varies around its mean value $\mathbb{E}[f(x)]$. Expanding out the square, we see that the variance can also be written in terms of the expectations of $f(x)$ and $f(x)^2$:

Exercise 2.8

$$\mathrm{var}[f] = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2. \tag{2.45}$$

In particular, we can consider the variance of the variable x itself, which is given by

$$\mathrm{var}[x] = \mathbb{E}[x^2] - \mathbb{E}[x]^2. \tag{2.46}$$

For two random variables x and y, the *covariance* measures the extent to which the two variables vary together and is defined by

$$\begin{aligned} \mathrm{cov}[x, y] &= \mathbb{E}_{x,y}\left[\{x - \mathbb{E}[x]\} \{y - \mathbb{E}[y]\} \right] \\ &= \mathbb{E}_{x,y}[xy] - \mathbb{E}[x]\mathbb{E}[y]. \end{aligned} \tag{2.47}$$