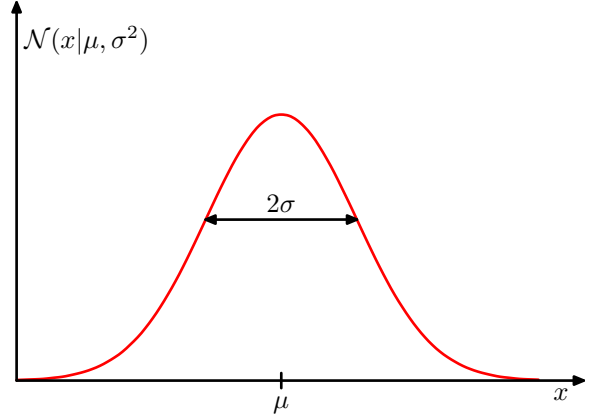


**Figure 2.8** Plot of a Gaussian distribution for a single continuous variable  $x$  showing the mean  $\mu$  and the standard deviation  $\sigma$ .



*Exercise 2.9*

If  $x$  and  $y$  are independent, then their covariance equals zero.

For two vectors  $\mathbf{x}$  and  $\mathbf{y}$ , their covariance is a matrix given by

$$\begin{aligned} \text{cov}[\mathbf{x}, \mathbf{y}] &= \mathbb{E}_{\mathbf{x}, \mathbf{y}} [\{\mathbf{x} - \mathbb{E}[\mathbf{x}]\} \{\mathbf{y}^T - \mathbb{E}[\mathbf{y}^T]\}] \\ &= \mathbb{E}_{\mathbf{x}, \mathbf{y}} [\mathbf{x} \mathbf{y}^T] - \mathbb{E}[\mathbf{x}] \mathbb{E}[\mathbf{y}^T]. \end{aligned} \quad (2.48)$$

If we consider the covariance of the components of a vector  $\mathbf{x}$  with each other, then we use a slightly simpler notation  $\text{cov}[\mathbf{x}] \equiv \text{cov}[\mathbf{x}, \mathbf{x}]$ .

## 2.3. The Gaussian Distribution

One of the most important probability distributions for continuous variables is called the *normal* or *Gaussian* distribution, and we will make extensive use of this distribution throughout the rest of the book. For a single real-valued variable  $x$ , the Gaussian distribution is defined by

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{1}{2\sigma^2} (x - \mu)^2 \right\}, \quad (2.49)$$

which represents a probability density over  $x$  governed by two parameters:  $\mu$ , called the *mean*, and  $\sigma^2$ , called the *variance*. The square root of the variance, given by  $\sigma$ , is called the *standard deviation*, and the reciprocal of the variance, written as  $\beta = 1/\sigma^2$ , is called the *precision*. We will see the motivation for this terminology shortly. [Figure 2.8](#) shows a plot of the Gaussian distribution. Although the form of the Gaussian distribution might seem arbitrary, we will see later that it arises naturally from the concept of maximum entropy and from the perspective of the central limit theorem.

From (2.49) we see that the Gaussian distribution satisfies

$$\mathcal{N}(x|\mu, \sigma^2) > 0. \quad (2.50)$$