2 Review of Basic Mathematical Concepts and Introduction to R

This chapter is a brief and informal review of some basic concepts from calculus and probability and an introduction to basic calculations using R. If you are not familiar with mathematical notions, you may also have to read introductory textbooks on calculus and statistics.

2.1 VARIABLES

A **variable** represents a varying quantity, for example, time t. It can take one of many values; for example, time $t = 2$ days denotes that t takes the value 2 in units of days. We will work mostly with variables that can take **real** values, for example, $t = 1.533$. Also, we define an interval of possible values; for example, t could take values in the interval from zero to 10 days, including the interval ends 0 and 10. It is common to denote a close interval like this by [0, 10]. We refer to this type of variable as **continuous**.

Exercise 2.1

Use a variable X to denote human population on Earth. Explain why it varies and give an example of a value.

2.2 FUNCTIONS

A **function** is a map that assigns a value of a dependent variable to each value of an independent variable. For example, $X(t)$ is a function X of t, where X is the **dependent** variable and t is the **independent** variable. For a more specific example, take

$$X(t) = at \tag{2.1}$$

which is a function expressing that variable X is proportional to the variable t, where the constant of proportionality is the coefficient a. This is an example of a **linear** function (Figure 2.1).

As another example, consider the function

$$X(t) = at^b \tag{2.2}$$

where the independent variable has been raised to a power b. When $b = 2$, this is a quadratic equation and it is called a **parabola**. We obtain a **nonlinear** function when the exponent is different from one, $b \neq 1$; see Figure 2.2.

In this book, we will use a wide variety of nonlinear functions. An example is given as

$$X(t) = \frac{a}{t + b} \tag{2.3}$$

which corresponds to a **hyperbolic** form. Using Equation 2.3 with $a = 2$ and $b = 1$, we obtain the graph shown in Figure 2.3.

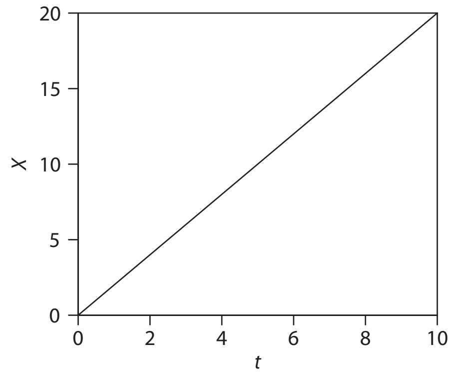

FIGURE 2.1 A linear function.

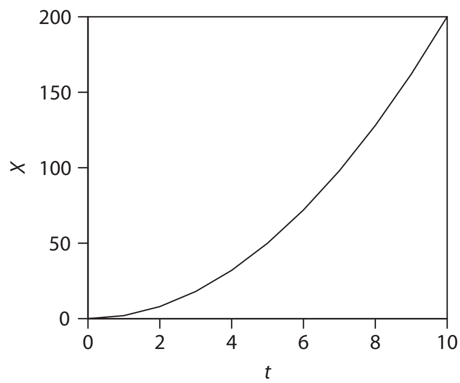

FIGURE 2.2 A parabolic function.

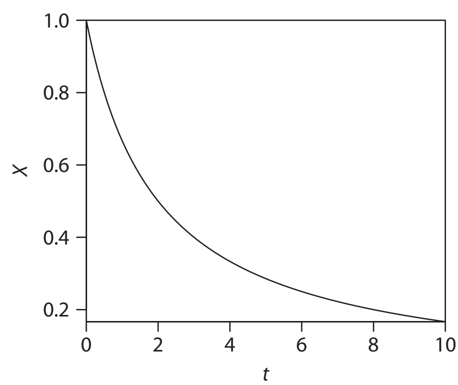

FIGURE 2.3 A hyperbolic function.

Another example is given as

$$X(t) = \exp(at) = e^{at} \tag{2.4}$$

which corresponds to an **exponential** function. In this last equation, the number e = 2.71828 is the base of natural **logarithms**. Recall that a logarithm is the power to which a base is raised to obtain a given number. For example, using base 10, log(100) = 2 or $10^2 = 100$. We use the exponential function often in this book to express that variable X has a rate of change that increases or decreases linearly with X. We will cover the exponential function in detail in the next chapter.

2.3 DERIVATIVES AND OPTIMIZATION

A concept often used in environmental models and employed throughout this book is the **derivative** of a function. For example, the derivative of $X(t)$ with respect to time t is denoted by

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} \tag{2.5}$$

This derivative represents a **rate of change** of X with time t and is equal to the **gradient** or **slope** of X with respect to t. This assumes that X varies continuously along t. You can think of a derivative as a ratio of very small change of two variables, for example, a very small change ΔX of X divided by a very small change Δt of t. The derivative is approximately equal to the slope obtained from the ratio $\Delta X/\Delta t$ (Figure 2.4). Therefore,

$$\frac{\mathrm{d}X}{\mathrm{d}t} \sim \frac{X}{t} \tag{2.6}$$

when deltas ΔX and Δt are very small or **infinitesimal** and can be referred to as **differentials** $\mathrm{d}X$ and $\mathrm{d}t$ of X and t.

There are rules to calculate the derivative of a function. The simplest ones are for polynomials or sums of power functions, like the one in Equation 2.2. The derivative is the product of the exponent and the variable raised to the power minus one. So the derivative of X with respect to t for the X given in Equation 2.2 is

$$\frac{\mathrm{d}X}{\mathrm{d}t} = abt^{b-1} \tag{2.7}$$

Exercise 2.2

Assume that $a = 1$ and $b = 2$. Evaluate the derivative of X with respect to t for the X given in Equation 2.2 and plot it. Is the derivative a constant with respect to t? Is the derivative a linear function with respect to t?

One application of derivatives is finding an **optimum** (minimum or maximum) of a function with respect to a variable. At an optimum, the derivative takes a value of zero. For example, the quadratic function or the parabola

$$X(t) = a + (t - b)^2 \tag{2.8}$$

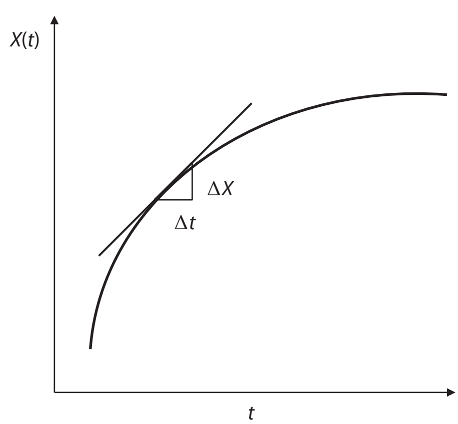

FIGURE 2.4 Derivative as the slope of the variable X as a function of time, t.

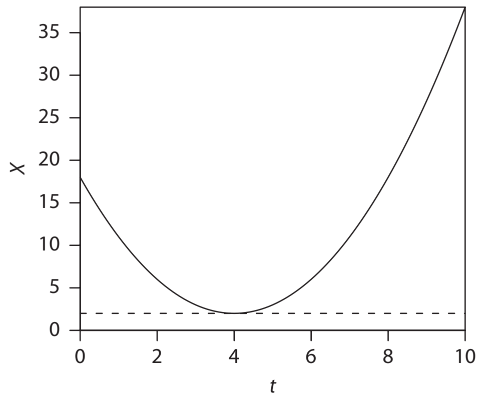

FIGURE 2.5 The derivative is zero at an optimum.

has a derivative

$$dX/dt = 2(t - b) \tag{2.9}$$

At an optimum, this derivative must be zero, and this will happen when the term in parentheses becomes zero, which occurs when $t = b$. Therefore, an optimum occurs at this value of t. Substituting in Equation 2.8, we see that the function takes the value

$$X(b) = a + (b - b)^2 = a$$

For $a = 2$ and $b = 4$, we can see that this optimum is a minimum with a value $a = 2$ and it occurs at $b = 4$ (Figure 2.5).

Exercise 2.3

Plot Equation 2.8 when $b = -4$ and $a = 3$. Find the values of the function and its derivative at $t = b$.

2.4 INTEGRALS AND AREA

You can think of the **integral** of a function as the inverse of the derivative and as the area under the curve representing the function within a defined interval. The integral is denoted by the integration symbol, ∫, applied to the function

$$\int_{t_1}^{t_2} X(t)dt \tag{2.10}$$

where the term after the integration symbol is called the integrand and t_1 and t_2 are called the limits of the integration interval. The differential dt indicates that the integration is performed over the variable t. The area is calculated as a summation of many small rectangles of width dx and height $X(t)$ when dx is infinitesimal (Figure 2.6) (Carr, 1995).

There are rules to calculate integrals; some of the commonly used ones follow. The integral of the inverse of X is the logarithm:

$$\int \frac{1}{X} dX = \ln(X)$$

The integral of a power function such as that in Equation 2.2 is

$$\int at^b dt = \begin{cases} a\dfrac{t^{b+1}}{b+1} & \text{for } b \neq -1 \\ a\ln(t) & \text{for } b = -1 \end{cases}$$

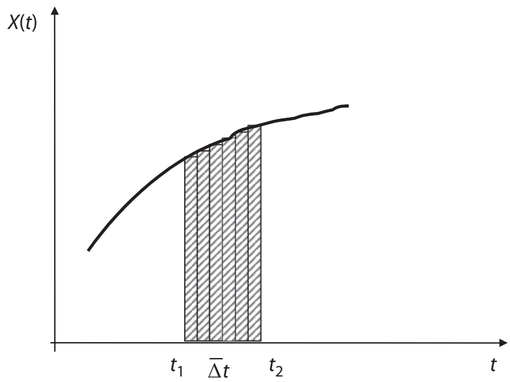

FIGURE 2.6 Integral as an area under the curve obtained by the summation of the area of many small rectangles of width Δt.

To evaluate for the given limits, one subtracts the integral evaluated at the lower limit from the integral evaluated at the upper limit. For example, if the limits for the integration of the term in Equation 2.2 were $t_1 = t_0 > 0$, $t_2 = t > t_0$, then

$$\int_{t_0}^{t} as^b \mathrm{d}s = \begin{cases} \dfrac{a}{b+1}\left(t^{b+1} - t_0^{\,b+1}\right) & \text{for } b \neq -1 \\ a\left(\ln(t) - \ln(t_0)\right) & \text{for } b = -1 \end{cases}$$

where we have used a dummy variable, s, for the integrand to distinguish it from the variable t used for the upper limit.

Exercise 2.4

Assume that $a = b = 1$ for the term in Equation 2.2. What is the area under the curve between $t = 2$ and $t = 6$? Illustrate the area using a graph like the one given in Figure 2.6.

2.5 ORDINARY DIFFERENTIAL EQUATIONS

Derivatives are used to build **differential equations** or **dynamic models** that state that the **rate of change** of X at any given time t, is a function $f(\cdot)$ of the value of X at that time t. These are **ordinary differential equations** (or **ODE** for brevity):

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f(X(t)) \tag{2.11}$$

When the function $f(\cdot)$ is linear, that is, the rate of change of a variable = coefficient × variable, we get a linear ODE. Using the derivative of X with respect to time t for the rate of change of X, we write

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = aX(t) \tag{2.12}$$

where the coefficient a is a **per unit** rate of change. We will discuss this type of linear ODE with more depth in the next chapter.

The solution of the ODE given in Equation 2.11 is a function, X, that satisfies the ODE. A solution of Equation 2.11 is found by separating the terms in X and t on different sides of the equation,

$$\frac{1}{f(X)}\,\mathrm{d}X = \mathrm{d}t \tag{2.13}$$

and by integrating both sides,

$$\int_{X(t_0)}^{X(t_1)} \frac{1}{f(X)} \mathrm{d}X = \int_{t_0}^{t_1} \mathrm{d}t \tag{2.14}$$

which requires the limits of integration for t and X. Here we define the initial time t_0 and the time t_1 at which we want to evaluate the solution. To evaluate the solution $X(t_1)$, we need to know the initial value $X(t_0)$ of X or the value of X at the initial time t_0. Depending on the complexity of $f(X)$, the ODE can be solved analytically using Equation 2.14. Often, we resort to a numerical calculation as explained in Chapter 4.

An extremely simple example will be the one where $f(X)$ is a constant a. In this case, the ODE is given as $\frac{\mathrm{d}X}{\mathrm{d}t} = a$. To solve it, apply Equation 2.14,

$$\int_{X(t_0)}^{X(t_1)} \frac{1}{a} \mathrm{d}X = \int_{t_0}^{t_1} \mathrm{d}t$$

which yields

$$\frac{1}{a}\big(X(t_1) - X(t_0)\big) = t_1 - t_0$$

and after solving for $X(t_1)$,

$$X(t_1) = a(t_1 - t_0) + X(t_0)$$

we obtain a straight line with slope a starting at the initial condition $X(t_0)$.

Exercise 2.5

Assume that $t_1 = t$ and $t_0 = 0$. Write the solution for the ODE of the example immediately above, using the initial condition $X(0) = 1$.

It is important to start getting familiar with the terminology used in modeling. We can do this at this early stage by recognizing how some of the terms apply to this simple model as follows:

- $\mathrm{d}X/\mathrm{d}t = aX$ is a dynamic **model**. In this case, it is given by an **ODE**.
- $X(t)$ is the **state variable** at time t. In this case, the state could also be called the **output**. In other cases, the output is a function of the state.
- a is a **parameter**.
- $X(t_0)$ is the **initial condition** or **initial state** (technically it is also a parameter).

2.6 FUNCTIONS OF SEVERAL VARIABLES

In many instances, a variable is dependent on more than just one independent variable. For example, the variable X can be a function of time t and position z:

$$X(t, z) = at + bz \tag{2.15}$$

In this case, we can look at two-dimensional (2-D) graphs to visualize the function. For example, with $a = 2$ and $b = 1$, Equation 2.15 yields the results shown in Figures 2.7 and 2.8. Figure 2.7 represents the results as isolines or contour lines (or lines of equal value of X) and Figure 2.8 uses an image representation.

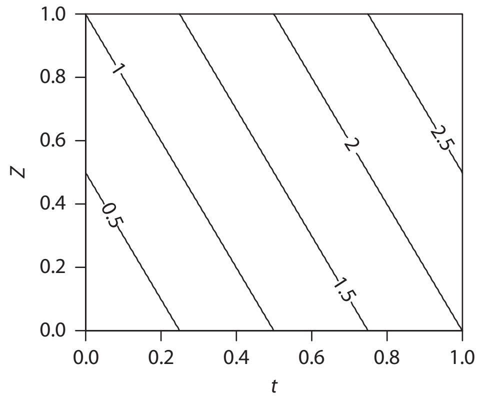

FIGURE 2.7 A linear function of two variables: an isoline or contour line view.

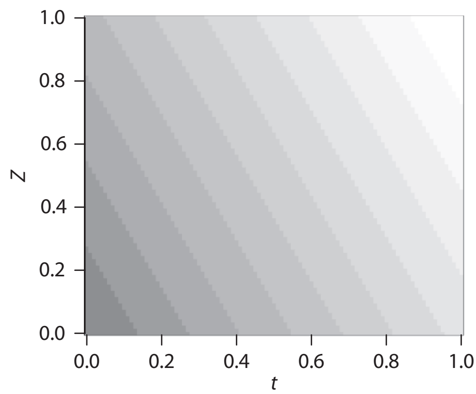

FIGURE 2.8 Function of two variables: image representation. The lighter the gray, the higher the value of X.

When we have more than one independent variable, we can take the derivative of the function with respect to each one of the variables. These are the **partial derivatives** of the function. In the example above, we can have two partial derivatives, one with respect to t and another with respect to z:

$$\frac{\partial X}{\partial t} = a \qquad \frac{\partial X}{\partial z} = b \tag{2.16}$$

Note that the symbol ∂/∂ is used for the partial derivative, which is different from the one used for ordinary derivatives, d/d.

One application of multivariable functions and partial derivatives is finding an **optimum** (minimum or maximum) of a function with respect to a set of parameters. At an optimum, all the partial derivatives take a value of zero. We will use this concept in the next chapter.

2.7 RANDOM VARIABLES AND DISTRIBUTIONS

A random variable (RV) is a rule or a map that associates a **probability** to each **event** in a sample space. The events can be defined from the intervals present in a range of real values; in this case, the values of an RV, X, are continuous in this range. We call this type of RV **continuous**. Consider the following example: measurements of concentration of a mineral (in parts per million [ppm]) at a given location can take values between 0 and 10,000 ppm. We may be interested in an event defined as that occurring when the measured concentration is in the interval 10–15 ppm.

A probability density function (pdf) $p(X)$ is based on the intervals; the probability of a value being in an infinitesimal interval of X between x and $x + \mathrm{d}x$ (Figure 2.9) is given by

$$p(x)\mathrm{d}x = P[x < X \le x + \mathrm{d}x] \tag{2.17}$$

where P denotes probability. Here, $p(x)$ is always positive or zero, that is, $p(x) \ge 0$.

The probability of a value being in an interval of X between a and b can be found using the integral

$$P[a < X \le b] = \int_a^b p(x)\mathrm{d}x \tag{2.18}$$

which is the area under the curve $p(x)$ for a given interval from a to b (Figure 2.10). When the interval is the whole range of values of X, then the value of the integral should be 1:

$$\int_{-\infty}^{+\infty} p(x)\mathrm{d}x = 1 \tag{2.19}$$

We have indicated the entire range of real values by selecting the limits from minus infinity ($-\infty$) to plus infinity ($+\infty$) or in practical terms from a very large negative value to a very large positive value.

Consider, for example, a uniform continuous RV. The density has the same value over the range $[a,b]$,

$$U_{a,b}(x) = \begin{matrix} \dfrac{1}{b-a} & \text{for} \quad a \le x \le b \\ 0 & \text{otherwise} \end{matrix} \tag{2.20}$$

as shown in Figure 2.11.

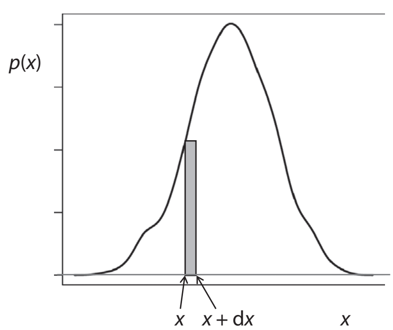

FIGURE 2.9 Probability density function of a continuous random variable. Probability is area under the curve between two values.

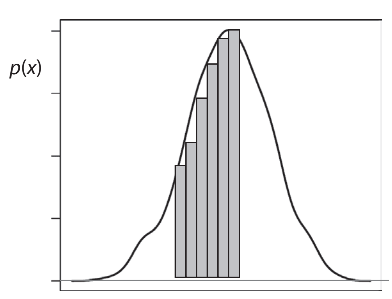

FIGURE 2.10 Probability of X having a value between a and b.

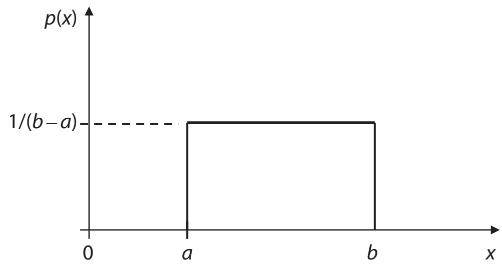

FIGURE 2.11 Probability density function of a uniform random variable.

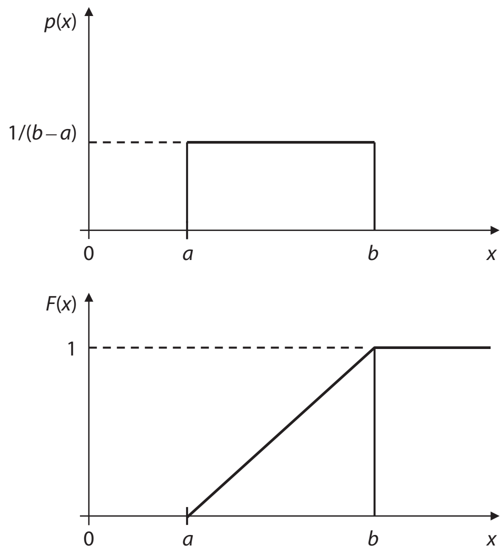

FIGURE 2.12 Probability density function and cumulative density function for a uniform random variable. Integration of a constant yields a linear increase (ramp function).

The "cumulative" distribution function (cdf) at a given value is defined by "accumulating" all probabilities up to that value:

$$F(x) = P[X \leq x] = \int_{-\infty}^{x} p(s)\mathrm{d}s \tag{2.21}$$

Note that the value at which we evaluate the cdf is the upper limit of the integral. Variable s is a dummy variable to avoid confusion with x. The cdf $F(x)$ at a value x is the area under the density curve up to that value. The value of the cdf for the largest value of X is its largest value and should be equal to 1. For example, for the uniform continuous RV, $U_{a,b}(x)$ is a ramp with slope $1/(b - a)$; see Figure 2.12.

2.8 RANDOM VARIABLES AND MOMENTS

The **first moment** of X is the **expected value** of X denoted by operator $E[\ ]$ applied to X, that is $E[X]$. This is the same as the **mean** of X. When X is continuous,

$${}_{X} = E[X] = \int_{-\infty}^{+\infty} xp(x)\mathrm{d}x \tag{2.22}$$

Note that the integration is over all values of X.

As an example, consider an RV uniformly distributed in [0,1]. In this case, $b = 1$ and $a = 0$. We know that $p(x) = 1/(b - a) = 1$:

$$ {}_X = E[X] = \int_0^1 x\mathrm{d}x = \frac{1}{2}x^2\Big|_0^1 = 1/2 \tag{2.23}$$

The expected value or mean is a theoretical concept. To calculate it, we need the pdf of the RV. The mean is not the same as the **statistic** known as the **sample mean**, which is the arithmetic average of n data values x_i comprising a sample:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.24}$$

Note that the mean (first moment) is denoted with the Greek letter μ, whereas the sample mean or average is denoted with a bar on top of X, that is to say, $\bar{X}$.

The second **central** (i.e., with respect to the mean) moment is the **variance** or the expected value of the square of the difference with respect to the mean:

$$\sigma_X^2 = E[(X - {}_X)^2] \tag{2.25}$$

If X is continuous, then

$$\sigma_X^2 = E[(X - {}_X)^2] = \int_{-\infty}^{+\infty} (x - {}_X)^2 p(x)\mathrm{d}x \tag{2.26}$$

The expectation $E[.]$ is calculated by an integral when the RV is continuous. The variance is a theoretical concept. To calculate it, we need the pdf of the RV. The **standard deviation** is the square root of this variance:

$$\sigma_X = \sqrt{\sigma_X^2}$$

From the definition of variance given in Equation 2.25, we can derive a more practical expression by substituting ${}_X = E[X]$ and expanding the square of a sum to obtain

$$\sigma_X^2 = E[(X - E(X))^2] = E[X^2 - 2XE(X) + E(X)^2]$$

Then, take the expected value of each term and use the fact that the expected value of a constant is the same constant to obtain

$$\sigma_X^2 = E[X^2] - 2E(X)E(X) + E(X)^2 = E[X^2] - 2E(X)^2 + E(X)^2$$

and finally

$$\sigma_X^2 = E[X^2] - E(X)^2 = E[X^2] - {}_X^2 \tag{2.27}$$

The variance or second central moment is not the same as the **statistic** known as the **sample variance**, which is the variability measured relative to the arithmetic average of n data values, x_i, comprising a sample:

$$\mathrm{var}(X) = s_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2 \tag{2.28}$$

This is the average of the square of the deviations from the sample mean. Alternatively,

$$\mathrm{var}(X) = s_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{X})^2 \tag{2.29}$$

where $n - 1$ is used to account for the fact that the sample mean was already estimated from the n values. We write s_X^2 to denote the sample variance to differentiate from the variance σ_X^2.

This equation can be converted to a more practical one by using Equation 2.24 in Equation 2.29 and doing algebra to obtain

$$\mathrm{var}(X) = s_X^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right] \tag{2.30}$$

This is easier to calculate because we can sum the squares of x_i and subtract the square of the sum of the x_i (Carr, 1995; Davis, 2002). Standard deviation is the square root of the variance. In this case, it refers to the standard deviation of a sample, which is different from the theoretical standard deviation.

The first and second central moments (that is, mean and variance) are also referred to as **parameters** of the RV (Carr, 1995; Davis, 2002). You have to be careful to avoid confusion with the term parameter as applied to an equation or a model. The mean and variance are different from the **statistics**, which are associated with the **sample** (Davis, 2002).

Another way of looking at this is to think of the pdf as a theoretical model expressing the underlying probability structure of the RV. These functions allow for the calculation of the moments. However, the statistics are calculated from the observed data and are used to **estimate** the moments.

2.9 SOME IMPORTANT RANDOM VARIABLES

The **uniform** RV has a pdf as given in Equation 2.20. The mean and variance are

$$\mu_X = \frac{b+a}{2} \qquad \sigma_X^2 = \frac{(b-a)^2}{12} \tag{2.31}$$

As an example, consider an RV uniformly distributed between a = 0, b = 1. The mean is 1/2 and the variance is 1/12.

The **normal**, or **Gaussian** RV, has a pdf given as

$$N_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \text{for} \quad -\infty < x < +\infty \tag{2.32}$$

with mean μ and variance σ^2. This is a symmetrical pdf, that is, the area under the curve left of the mean is the same as the area under the curve to the right of the mean. The area of the curve on both sides of the mean increases with distance: at one standard deviation on both sides ($\mu \pm \sigma$), the area is 0.68, at two standard deviations ($\mu \pm 2\sigma$), the area is 0.95, and at three standard deviations ($\mu \pm 3\sigma$), the area is 0.99 (Davis, 2002).

As another example, consider an RV normally distributed with mean = 1 and variance = 0.25. What is the probability of obtaining a value in between 0.5 and 1.5? The standard deviation is 0.5. This interval is one standard deviation away from the mean on each side. Therefore, the probability is 0.68.

A **standard normal** RV is a normal RV with a mean equal to zero ($\mu = 0$) and variance equal to one ($\sigma^2 = 1$). To obtain a standard normal RV from a normal RV, we subtract the mean and divide it by the standard deviation:

$$Z = \frac{X - \mu_X}{\sigma_X} \tag{2.33}$$

The new RV Z is standard normal. Its mean is 0 and variance is 1. All values to the left of the mean are negative ($z < 0$) and all values to the right of the mean are positive ($z > 0$). Note that z is scaled in units of standard deviation. Because the normal is symmetric, calculating the area under the standard pdf curve from $-\infty$ up to a value $-z_0$ (left of the mean) is the same as calculating the area under the curve from that value $+z_0$ to $+\infty$ (right of the mean); see Figure 2.13.

Exercise 2.6

Define an RV from the outcome of soil moisture measurements in the range of 20–100% in volume. Give an example of an event. Assuming that it can take values in [20,100] uniformly, plot the pdf and cdf. Calculate the mean and variance.

Exercise 2.7

At a site, the monthly air temperature is normally distributed. It averages to 20°C with a standard deviation of 4°C. What is the probability that the value of the air temperature in a given month exceeds 24°C? What is the probability that it is below 16°C or above 24°C?

The exponential pdf is given by

$$p(x) = a\exp(-ax) \tag{2.34}$$

and the cdf is given by

$$F(x) = 1 - \exp(-ax) \tag{2.35}$$

It has mean and variance equal to $_X = 1/a$ and $\sigma_X^2 = 1/a^2$, respectively.

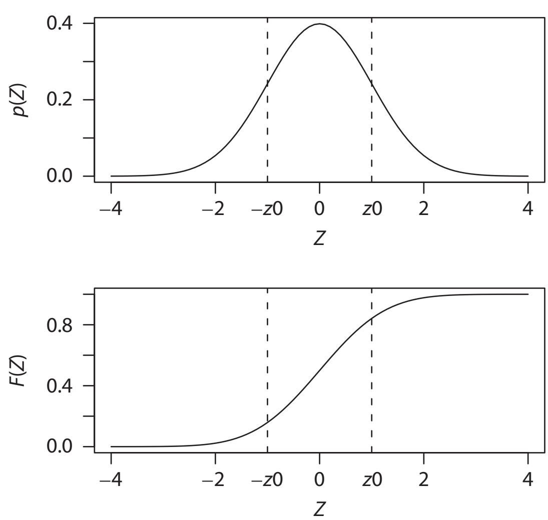

FIGURE 2.13 Standard normal pdf and cdf.

2.10 COVARIANCE AND CORRELATION

In many cases, we are interested in how variables relate to each other. The **bivariate** or simplest case is that of **two** RVs, X and Y. An important concept is the joint variation or the expected value of the product of the two variables, where each one is centered at the mean. This is called the **covariance** and can be written as

$$\mathrm{cov}(X,Y) = E[(X - \ {}_X)(Y - \ {}_Y)] \tag{2.36}$$

Please note this is a theoretical concept, since the expectation operator implies using the distribution of the product. Therefore, we require the joint pdf of X and Y.

A derived concept is the **correlation coefficient** obtained by scaling the covariance to values less than or equal to 1. The scaling factor is the product of the two independent standard deviations:

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{2.37}$$

Because this product is always larger than the expected value of the product, the ratio is always less than 1. This fact can also be seen by calculating the correlation coefficient for maximum covariance, which occurs when X and Y are identical:

$$\rho = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{cov}(X,X)}{\sigma_X \sigma_X} = \frac{\sigma_X^2}{\sigma_X^2} = 1 \tag{2.38}$$

Expanding Equation 2.36, and since the expectation of a constant is the same constant, we obtain

$$\begin{aligned} \mathrm{cov}(X,Y) &= E[(XY - Y \ {}_X - X \ {}_Y + \ {}_X \ {}_Y)] \\ &= E[XY] - E[Y] \ {}_X - E[X] \ {}_Y + \ {}_X \ {}_Y \\ &= E[XY] - \ {}_Y \ {}_X - \ {}_X \ {}_Y + \ {}_X \ {}_Y \\ &= E[XY] - \ {}_X \ {}_Y \end{aligned} \tag{2.39}$$

The same idea can be applied to a sample to obtain the **sample correlation coefficient**, r, where the covariance is a **sample covariance** and the denominator corresponds to the product of the standard deviations of the sample.

2.11 COMPUTER SESSION

2.11.1 Variables

When we use a continuous variable on a computer, we can only use a discrete approximation based on a set of values. Having more values in the interval improves the approximation. There are many ways to assign a set of values to a variable in R. For example, we can use a sequence from zero to 10 in steps of one:

```
> t <- seq(0,10,1)
```

Recall that the operator "<-" is used for assignment. Double-check that you have the newly created object by using `ls()`:

```
> ls()
[1] "t"
```

Object `t` is temporarily stored in **.Rdata** (workspace). We can double-check the contents of the object by typing its name:

```
> t
 [1]  0 1 2 3 4 5 6 7 8 9 10
```

We can see that this object is a one-dimensional (1-D) array. The number in brackets on the left-hand side is the position of the entry first listed in that row. For example, the entry in position 1 is zero. Object `t` should have 11 values, and this can be verified using the function `length`:

```
> length(t)
 [1]  11
```

We can obtain a better approximation to a continuous variable by changing the step size to 0.1:

```
> t <- seq(0,10,0.1)
> t
  [1]  0.0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9  1.0  1.1  1.2  1.3  1.4  1.5
 [17]  1.6  1.7  1.8  1.9  2.0  2.1  2.2  2.3  2.4  2.5  2.6  2.7  2.8  2.9  3.0  3.1
 [33]  3.2  3.3  3.4  3.5  3.6  3.7  3.8  3.9  4.0  4.1  4.2  4.3  4.4  4.5  4.6  4.7
 [49]  4.8  4.9  5.0  5.1  5.2  5.3  5.4  5.5  5.6  5.7  5.8  5.9  6.0  6.1  6.2  6.3
 [65]  6.4  6.5  6.6  6.7  6.8  6.9  7.0  7.1  7.2  7.3  7.4  7.5  7.6  7.7  7.8  7.9
 [81]  8.0  8.1  8.2  8.3  8.4  8.5  8.6  8.7  8.8  8.9  9.0  9.1  9.2  9.3  9.4  9.5
 [97]  9.6  9.7  9.8  9.9 10.0
>
```

Important tip: you can recall previously typed commands using the up arrow key. For example, you can use the up arrow key and recall the line already used to apply length:

```
> length(t)
[1]  101
```

We can refer to specific entries of an array using brackets or square braces. For example, the entry in position 24 of `t` is 2.3:

```
> t[24]
[1]  2.3
```

A colon symbol (“:”) declares a sequence of entries. For example, the first 10 positions of `t`, can be listed as follows:

```
> t[1:10]
[1]  0.0  0.1  0.2  0.3  0.4  0.5  0.6  0.7  0.8  0.9
```

To remove entries, use the minus sign; for example, use the following to remove the first entry of `t`:

```
> t[-1]
[1]  0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6
…. etc
```

As we can see, we have removed entry 0.0. In addition, we can confirm by using the length:

```
> length(  t[-1])
[1]  100
```

A blank inside the brackets addresses all the entries. For example,

```
> t[]
[1] 0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5
… etc
```

This is the same as if we would have used just `t`.

2.11.2 Plots

As an example, let us use t in the interval [0,10] and $a = 2$ in Equation 2.1. To evaluate a function in R, first define the independent variable, next give a value to the parameter, and then apply the function. For example, define t as a sequence from 0 to 10 in steps of 1 and then apply a multiplication:

```
t <- seq(0,10,1)
a <- 2
X <-a*t
```

All values of array t are multiplied by a, and therefore, the new array X contains 11 values of variable X:

```
> X
[1] 0 2 4 6 8 10 12 14 16 18 20
```

A function can be represented as a graph in the x-y plane, where the x-axis or horizontal axis is used for the independent variable and the y-axis or vertical axis is used for the dependent variable. To visualize the result, we can run a simple plot function:

```
>plot(t,X)
```

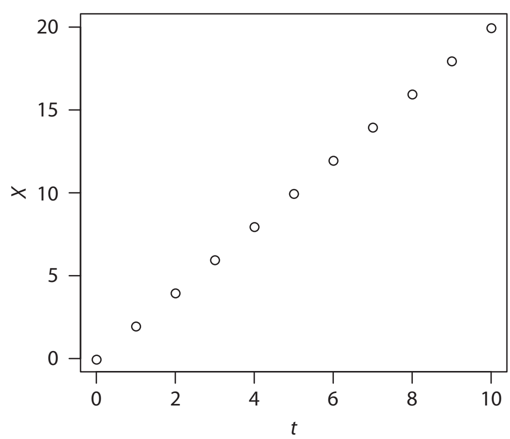

FIGURE 2.14 A plot of a linear function.

A graphics function will send the information to the active graphics window or create a new graphics window if needed. We can see that we obtain Figure 2.14, where the graph shows a circle symbol for each pair of values t, `X`. We can run the plot function using a line graph, obtained with the argument `type = "l"`. Here "l" is a letter character for "line." Do not confuse it with the number one:

```
>plot(t,X, type="l")
```

gives Figure 2.15.

Using the arguments `xaxs = "i"` and `yaxs = "i"`, the plot function skips the margins inside the plot area:

```
>plot(t,X, type="l", xaxs="i", yaxs="i")
```

as shown in Figure 2.1. Instead of including these arguments in the call to **plot**, it is often practical to first set graphics parameters using `par` before calling the `plot` function. This way all subsequent calls to **plot** will use these settings while the graphics device remains open. For example, the following will produce both plots with internal style "i" for both axes:

```
>par(xaxs="i",yaxs="i")
>plot(t,X)
>plot(t,X, type="l")
```

The contents of a graphics window can be saved in image format using **File|Save as**. For example, it can be saved as JPEG in one of three quality values, or TIFF, or PDF. You could also simply copy or save with a right-click on the graphics window or **File|Copy to clipboard** and then paste the contents into an application.

Exercise 2.8

Draw line graphs of the function given in Equation 2.1 for several values of coefficient *a*. Use *a* = 0.1, 1, 10. Produce three graphs.

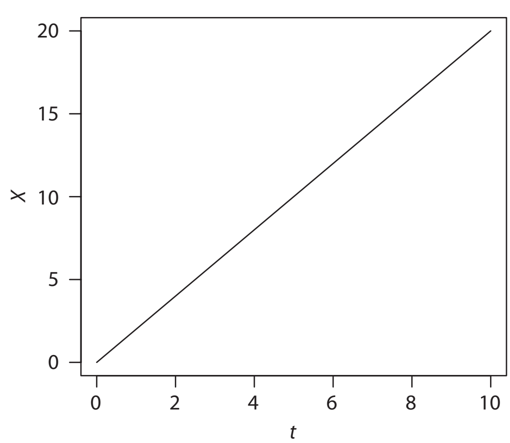

FIGURE 2.15 A line graph plot of a linear function.

Figure 2.2 can be produced with the following script where we used $a = 2$, $b = 2$. A semicolon is used to separate statements in the same line, and the symbol ^ is used for the power operator:

```
t <- seq(0,10,1)
a <- 2; b <-2
X <- a*t^b
plot(t,X, type="l")
```

Exercise 2.9

Plot Equation 2.2 using $a = 2$ and three different values of b: 0.1, 1, and 3.

With a script just like the previous one, we can produce Figure 2.3 except that we use `X <- a/(t+b)` and increment the number of values in the set for X by using time step of 0.1 so that the graph looks more continuous.

2.11.3 Defining Functions in R

We will use the hyperbolic function to introduce how to define functions in R. The following line assigns f as a function with arguments t, a, and b:

```
f <- function(t, a, b) X <- a/(t+b)
```

Now, to call the function, first give values to the arguments and use the name of the function:

```
t <- seq(0,10,0.1)
a <- 2; b <-2
X <- f(t,a,b)
plot(t,X, type="l")
```

Once we store a function in the workspace, we can write a call to it with proper arguments. It is available whenever we need it as long as we save the workspace before we exit the session. You can verify that the function, f, is stored in your workspace using `ls()`.

Exercise 2.10

Generate values for a function $y = ax + b$, $a = 0.1$, $b = 0.1$. Plot y for values of x in 0 to 1.

2.11.4 Exponential

For example, first define the independent variable as a sequence from 0 to 10 in steps of 0.1, then give a value to the rate coefficient, calculate variable `X` for all t using function `exp`, and plot:

```
t <- seq(0,10,0.1)
r<-0.1
X <- exp(r*t)
plot(t,X, type="l")
```

See the result in Figure 2.16. The limits of the x and y axes of the graph can be changed using `xlim=c(,)` and `ylim=c(,)`, where c(,) denotes a 1-D array with two entries. For example, `xlim=c(0,5)` has two elements, a minimum of 0 and a maximum of 5, and `ylim=c(0,2)` has two elements, a minimum of 0 and a maximum of 2.

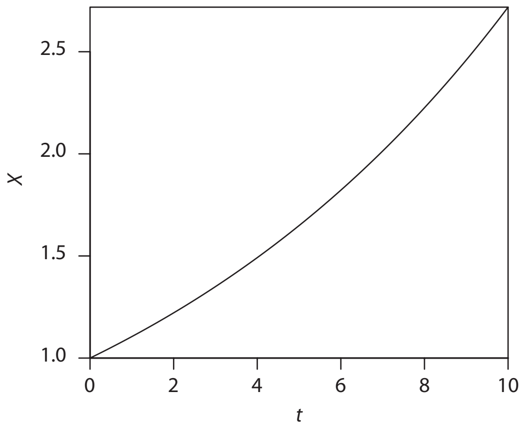

FIGURE 2.16 An exponential function with t taking values in [0, 10] and $r = 0.1$.

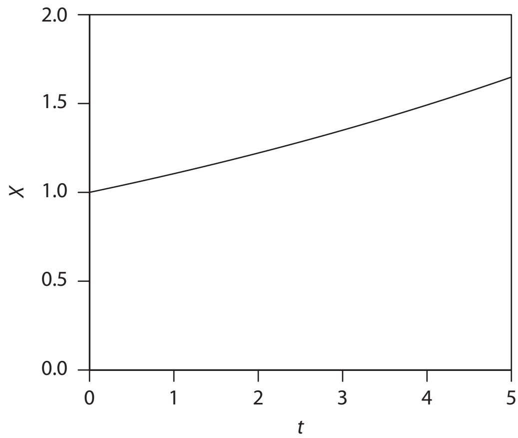

FIGURE 2.17 Changing the limits of the x and y axes of the graph.

```
>plot(t,X, type="l", xlim=c(0,5), ylim=c(0,2))
```

See Figure 2.17.

Exercise 2.11

Generate values for an exponential function with r = −0.1. Plot X for values of t in 0 to 10. Then limit the *t*-axis to interval [0,1].

2.11.5 Derivatives

A computer cannot take the exact derivative but it can calculate an approximation using Equation 2.6. Using R, we have the function `diff`, which can be used to calculate ΔX when X is given by an array. For example, assume we calculate `X` as a linear function as in the previous section:

```
> dt <- 0.1 # set time step
> t <- seq(0,10,dt)
> a <- 2; X <-a*t  # parameter and function
```

Now we take the difference and approximate the derivative by the ratio $\Delta X/\Delta t$:

```
> dX <- diff(X)
> dX.dt <- dX/dt
> dX.dt
 [1] 2 2 2 2 2 2 2 2 2 2 2 2 2  etc
```

We see that the derivative is a constant 2, which is the constant slope of the linear function. Note that the resulting array is shorter than the original because the first difference is the second entry minus the first entry of X.

```
> length(dX)
[1] 100
> length(X)
[1] 101
```

Exercise 2.12

Assume that $a = 1$ and $b = 2$. Calculate the derivative of X with respect to t for X as given in Equation 2.2 using **diff** and plot it. Compare this with Exercise 2.2. Hint: The first entry of t should be removed so that dX.dt and t have the same length and are compatible for the function plot.

We can evaluate the minimum or maximum of an array or a matrix using **min** and **max** functions. For example, define a function f for Equation 2.8 and evaluate it at a = 2, b = 4.

```
f <- function(t,a,b) {
  X <- a +(t-b)^2
  }
t <- seq(0,10,0.1)
X <- f(a=2,b=4,t)
```

We can use min to find the minimum and which to find the entry number at which the minimum occurs:

```
>min(X); t[which(X==min(X))]
[1] 2
[1] 4
```

2.11.6 Reading Data Files

As an example, let us work with a data set of 100 numbers drawn from a normal distribution of mean = 40 and standard deviation = 10 that are stored in a file **chp2\test100.txt**. Use a text editor (e.g., Notepad, Vim) to look at the file. Again, it is convenient to configure the list of files to show all the file extensions. Otherwise, you will not see the **.txt** part of the filename.

Because the working directory is **seem**, the path to the file is relative to this folder, for example, in this case **chp2\test100.txt**. Therefore, you could use this name to scan the file. Use forward slash "/" to separate folder and filename. Next, create an object, X, by reading or scanning this data file:

```
> X <- scan("chp2/test100.txt")
```

Values read or scanned from the file are assigned to object `X`. Double-check that you have the newly created object by using `ls()`:

```
> ls()
[1] "X"
```

Note: Object `X` will be stored in **seem\.Rdata** (i.e., in the workspace), but file **test100.txt** resides in **seem\chp2**. Double-check the object contents by typing its name.

```
> X
  [1] 48 38 44 41 56 45 39 43 38 57 42 31 40 56 42 56 42 46 35 40 30 49 36 28 55
 [26] 29 40 53 49 45 32 35 38 38 26 38 26 49 45 30 40 38 38 36 45 41 42 35 35 25
 [51] 44 39 42 23 44 42 52 55 46 44 36 26 42 31 44 49 32 39 42 41 45 50 39 55 48
 [76] 49 26 50 46 56 31 54 26 29 32 34 40 53 37 27 45 37 34 32 33 35 50 37 74 44
```

We can see that this object is an array. As we mentioned already, the number in brackets on the left-hand side is the position of the entry first listed in that row. For example, the entry in position 26 is 29 and the entry in position 51 is 44. Since this object is a 1-D array, we can check the size of this object by using command `length`:

```
> length(X)
[1] 100
```

Plot a histogram by using the function `hist` applied to a single variable (univariate or 1-D) object. Example:

```
>hist(X)
```

On the window corresponding to the graphics device, you will see a histogram like that shown in Figure 2.18. You could use a probability density scale (the bar height times the width would be a probability, i.e., from 0 to 1) on the vertical axis by using `hist` with the option `freq=F`:

```
>hist(X, freq=F)
```

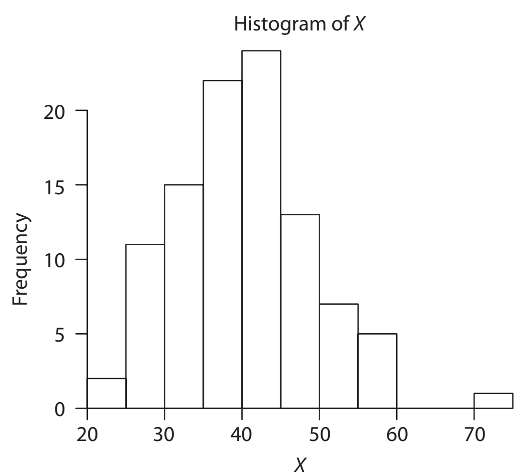

FIGURE 2.18 Histogram in frequency units.

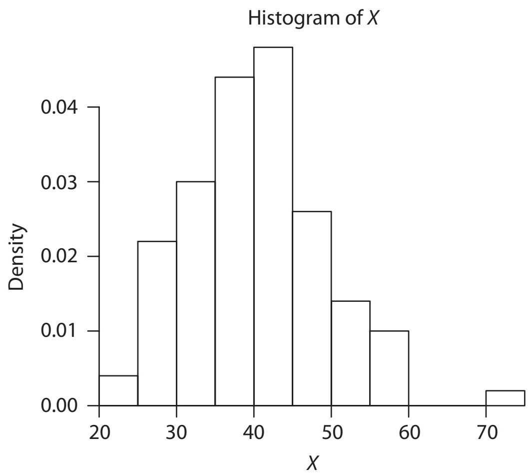

FIGURE 2.19 Histogram using probability units.

R handles many options in a similar manner: it gives a logical variable the value `T` (for **T**rue) and `F` (for **F**alse by default). Here, by default, `T` corresponds to the bar height that is equal to the count. The result in this case is given in Figure 2.19.

We can calculate sample mean, variance, and standard deviation using the functions `mean`, `var`, and `sd`.

```
> round(mean(X),2);round(var(X),2); round(sd(X),2)
[1] 40.86
[1] 81.62
[1]  9.034
```

The sample mean and standard deviation approach the population mean and standard deviations of 40 and 10, respectively.

Part II

One-Dimensional Models and Fundamentals of Modeling Methodology