

Discrete Difference Equations

from Mathematics for the Life Sciences, book by Bodine, Lenhart and Gross

These notes discuss how we can use sequences from data or from theory to determine how biologically relevant variables change. One of the main objectives is to learn how to use mathematical formulae to describe a sequence and what the long-term implications are for the variable being measured. One example is using your sequence of heights measured from birth to age 10 to predict what your adult height would be. Note that we are using height as an example rather than weight because adult height is much more constrained, and does not change much over adulthood, as compared to body weight (you might consider why this is so).

EXAMPLE: Bird Count

Every Christmas the Audubon Society invites birders across the United States to participate in the Christmas Bird Count. Data for various birds have been collected for over 100 years. The data is posted on the Audubon Society's website. Suppose we form a sequence using the data collected on red cardinals each year. The count the first year would be a_1 , the second year would be a_2 , the third year would be a_3 , and so on.

Year	Count	Year	Count	Year	Count	Year	Count
1959	2206	1972	3696	1985	5359	1998	5439
1960	2297	1973	4989	1986	4321	1999	4367
1961	2650	1974	3779	1987	5044	2000	6045
1962	2277	1975	4552	1988	3092	2001	4632
1963	2242	1976	3872	1989	5388	2002	6974
1964	2213	1977	4049	1990	4079	2003	4528
1965	2567	1978	4037	1991	4416	2004	6875
1966	3152	1979	3475	1992	4828	2005	5154
1967	2186	1980	4448	1993	4291	2006	6631
1968	2998	1981	3660	1994	4861	2007	7051
1969	2628	1982	5141	1995	4662	2008	4882
1970	3450	1983	4890	1996	4827	2009	6896
1971	2829	1984	3500	1997	4377	2010	6190
						2011	6739

Table 1: Data collected by Tennessee birders for the Audubon Christmas Bird Count. Data reflects the total count of Northern Cardinals sighted in Tennessee.

Cardinals: Birders in Tennessee have reported seeing the Northern Cardinal (*Cardinalis cardinalis*) since 1959. The numbers reported are given in the Table 1. We could make a sequence out of this data. If a_n represents the count at year n , where $n = year - 1959$ then,

$$a_0 = 2206, a_1 = 2297, a_2 = 2650, \dots, a_{50} = 6896, a_{51} = 6190, a_{52} = 6739.$$

This particular sequence has 53 terms in the sequence.

A real sequence is a function $f : \mathbb{N} \rightarrow \mathbb{R}$ with domain being the natural numbers \mathbb{N} and the range being contained in the real numbers \mathbb{R} . We write $a_n = f(n)$ and call a_n the n^{th} term in the sequence.

Some real sequences can be expressed by an algebraic formulation.

EXAMPLE: Sequence

Suppose we defined the sequence $a_n = f(n) = (-1)^n \frac{2n}{n+1}$. Find the first 5 terms of this sequence.

Solution: We can use the function definition of the sequence to find each term.

$$f(1) = (-1) \frac{2 \cdot 1}{1 + 1} = -1$$

$$f(2) = (-1)^2 \frac{2 \cdot 2}{2 + 1} = \frac{4}{3}$$

$$f(3) = (-1)^3 \frac{2 \cdot 3}{3 + 1} = -\frac{6}{4} = -\frac{3}{2}$$

$$f(4) = (-1)^4 \frac{2 \cdot 4}{4 + 1} = \frac{8}{5}$$

$$f(5) = (-1)^5 \frac{2 \cdot 5}{5 + 1} = -\frac{10}{6} = -\frac{5}{3}$$

Limit of a Sequence

Our objective is to point out how we can use mathematical descriptions of sequences to analyze the behavior of simple models.

Suppose we have a sequence $\{a_n\}$ whose terms are getting closer and closer to L as n gets larger and larger. Then we say that the limit of the sequence $\{a_n\}$ is L . Formally, we write

$$\lim_{n \rightarrow \infty} a_n = L$$

if the the sequence $\{a_n\}$ converges to the limit L (a finite number). If the sequence $\{a_n\}$ increases toward infinity or decreases toward negative infinity as $n \rightarrow \infty$, then we say that the sequence $\{a_n\}$ does not have a limit. Thus, note that not all sequences have limits.

Discrete Difference Equations

Suppose we have a population that doubles each year. We could construct a sequence, where x_n represents the size of the population at each time step. Since the population is doubling each year, $x_{n+1} = 2x_n$, i.e. the population at time step $n + 1$ is twice what it was at time step n , using yearly time steps.

We will study equations showing biological quantities that changes over time. For example, the Fibonacci sequence,

$$1, 1, 2, 3, 5, 8 \dots$$

(originally involving the births in generations of rabbit populations), is given by:

$$x_0 = 1, x_1 = 1$$

and then for $n = 1, 2, \dots$

$$x_{n+1} = x_n + x_{n-1}.$$

Note that x_{n+1} is a function of the two previous terms in the sequence, which means that this equation for x_{n+1} is called a second order difference equation. Note that the equation above, $x_{n+1} = 2x_n$, is called a first order difference equation since each new term in the sequence only depends on the previous term.

If the x_{n+1} term can be written solely as a function of the x_n term, i.e.

$$x_{n+1} = f(x_n),$$

then we say that this sequence is built from a **first order difference equation or recursive equation**. The term ‘first order’ meaning that to find x_{n+1} , you only need to use the previous value, x_n . Notice, that if we know x_0 (the initial value in the sequence) then we can determine all the terms in the sequence, coming from a first difference equation.

EXAMPLE: Doves

A population of doves increases by 3% each year. Let x_n be the size of the population at year n . Then $x_{n+1} = x_n + .03x_n = 1.03x_n$. Thus, the first order difference equation that describes the population is

$$x_{n+1} = 1.03x_n.$$

Notice that if we know x_0 , then

$$\begin{aligned} x_1 &= 1.03x_0 \\ x_2 &= 1.03x_1 = 1.03(1.03x_0) = (1.03)^2x_0 \\ &\vdots \\ x_n &= (1.03)^nx_0. \end{aligned}$$

Geometric and Arithmetic Sequences

The sequence in the dove example above is called a geometric sequence since its population changes by a multiplication factor each time step.

A **geometric sequence** is defined by

$$x_{n+1} = rx_n$$

where r is a fixed real number. Notice, if we know x_0 , then

$$\begin{aligned} x_1 &= rx_0 \\ x_2 &= rx_1 = r(rx_0) = r^2x_0 \\ x_3 &= rx_2 = r(r^2x_0) = r^3x_0 \\ &\vdots \\ x_n &= r^nx_0 \end{aligned}$$

The general solution to difference equation represents the x_n in terms of x_0 , n and other given constants. For a geometric sequence, the general solution is is

$$x_n = r^n x_0.$$

When $0 < r < 1$, the sequence decays to zero. It has a limit of zero, meaning as n get large ($n \rightarrow \infty$), $x_n \rightarrow 0$. For $r > 1$, the terms of the sequence increase exponentially.

EXAMPLE: Hares

A population of wild hares increases by 13% each year. Currently, there are 200 hares. If x_n is the number of hares in the population at the end of year n , find

1. the difference equation relating x_{n+1} to x_n ,
2. the general solution to the difference equation found in (a), and
3. the number of hares in the population at the end of six years from now.

Solution: Notice that $x_0 = 200$.

1. Since the population increases by 13% each year,

$$x_{n+1} = \underbrace{x_n}_{\text{from year before}} + \underbrace{0.13x_n}_{\text{increase}} = 1.13x_n.$$

2. Using the notation of the definition of a geometric sequence, here $r = 1.13$, thus the general solution to the difference equation in (a) is

$$x_n = (1.13)^n x_0 = 200(1.13)^n.$$

3. For $n = 6$, $x_6 = 200(1.13)^6 \approx 416$. Thus, at the end of year six there are approximately 416 hares.

If a population increases by a fixed number d each time period, we say the sequence is an **arithmetic sequence**,

$$x_{n+1} = x_n + d.$$

Suppose we know the initial value x_0 , and then the general solution to an arithmetic sequence is:

$$\begin{aligned} x_1 &= x_0 + d \\ x_2 &= x_1 + d = (x_0 + d) + d = x_0 + 2d \\ &\vdots \\ x_n &= x_0 + nd. \end{aligned}$$

Linear Difference Equation with Constant Coefficients

A first order difference equation is linear if it takes the form

$$x_{n+1} = a_n x_n + b_n$$

where a_n and b_n are sequences of constants. In this section, we are interested in linear first order difference equations where $a_n = a$ and $b_n = b$ for all n . That is, we can write the difference equation as

$$x_{n+1} = ax_n + b.$$

Let us find the general solution to this difference equation.

First, notice that if $a = 1$, then we are in the case of the arithmetic sequence and we have already found the general solution,

$$x_n = x_0 + nb.$$

Second, if $b = 0$, we have a geometric sequence with the general solution $x_n = a^n x_0$.

If $a \neq 1$ and $b \neq 0$, then to find the general solutions, we use the following steps.

Finding the General Solution to $x_{n+1} = ax_n + b$ where $a \neq 1$

Step 1: We first solve the homogeneous difference equation, which is the difference equation without the constant b , that is $x_{n+1} = ax_n$. We already know the solution of this has the form $a^n c$ where c is some constant. Notice, we cannot say that $c = x_0$ here because a solution to the homogeneous equation is not necessarily a solution to the full equation.

Step 2: Next, we construct a particular solution p_n . We will assume the particular solution is a constant, that is $p_n = K$ for all n . We now need to determine K . For p_n to be a solution to the difference equation, it must satisfy

$$p_{n+1} = ap_n + b.$$

However, $p_n = K$, thus

$$\begin{aligned} K &= aK + b \\ K - aK &= b \\ (1 - a)K &= b \end{aligned}$$

$$K = \frac{b}{1 - a}.$$

Thus, our particular solution is $p_n = \frac{b}{1-a}$.

Step 3: Lastly, we form the general solution as the sum of the homogeneous and particular solutions. That is,

$$x_n = ca^n + \frac{b}{1 - a}.$$

Notice $x_0 = c + \frac{b}{1-a}$, thus $c = x_0 - \frac{b}{1-a}$.

To see why building a general solution in this way works, consider two solutions, s_n and w_n , to

$$x_{n+1} = ax_n + b.$$

Let $y_n = s_n - w_n$ and then y_n satisfies this difference equation,

$$y_{n+1} = s_{n+1} - q_{n+1} = (as_n + b) - (aw_n + b) = a(s_n - w_n) = ay_n.$$

We have y_n solving the homogeneous equation, giving $y_n = ca^n$ for a constant c . Any general solution can be written as $s_n = ca^n + w_n$.

Therefore, the general solution to the difference equation is

$$x_n = \left(x_0 - \frac{b}{1-a}\right)a^n + \frac{b}{1-a}.$$

Let us consider an example of a population that could be modeled by a linear difference equation with constant coefficients. What would each term in the difference equation represent?

$$\underbrace{x_{n+1}}_{\text{pop. @ } t = n + 1} = \underbrace{ax_n}_{\text{pop growth or decline}} - \underbrace{b}_{\text{fixed decrease}} \quad (1)$$

Notice that in Equation (1), the population grows or declines first (with the factor a) and then a fixed amount b is being removed. Thus, this model could represent a population that is being harvested by a fixed amount at each time step after growth or decline. If the constant b was being added in the equation, then the difference equation might model a population (after growth or decline) that was being augmented by a fixed amount each time step (think of a fish population being restocked each season).

About the **order of events**: If one harvests the population before the growth happens, then you would start with this equation:

$$x_{n+1} = a(x_n - b).$$

Furthermore, it may be advantageous to think of the constant a , representing population growth or decline, as the difference between the birth and death rates of the population. That is, if β is the proportion by which the population increases due to births at each time step, and δ is the proportion by which the population decreases due to deaths at each time step, then $a = 1 + \beta - \delta$. If $a > 1$ then the population is growing (in the absence of harvesting), if $a < 1$ then the population is decreasing (in the absence of harvesting), and if $a = 1$ then in the absence of harvesting the population will remain constant.

EXAMPLE: Fisheries

Consider a lake fish population whose yearly birth rate is 1.2, and yearly death rate is 0.7. Each year, fishing is allowed until 1200 fish are caught. Thereafter, fishing is banned. Currently, there are 12,230 fish in the lake.

1. Write a difference equation for the lake fish population and find the general solution.
2. How many fish are in the lake after 5 years?
3. If the resource managers of the lake wanted the population to remain constant each year, what level of harvesting should they allow?

Solution: Let x_n be the size of the fish population at the end of year n . Then $x_0 = 12,230$.

1. Since $\beta = 1.2$ and $\delta = 0.7$, then $a = 1 + \beta - \delta = 1 + 1.2 - 0.7 = 1.5$. Thus, $x_{n+1} = 1.5x_n - 1200$. The general solution is then

$$\begin{aligned} x_n &= \left(x_0 - \frac{-1200}{1 - 1.5} \right) (1.5)^n + \frac{-1200}{1 - 1.5} \\ &= (12,230 - 2400)(1.5)^n + 2400 \\ &= 9830(1.5)^n + 2400. \end{aligned}$$

2. For $n = 5$, $x_n = 9830(1.5)^5 + 2400 \approx 77,047$. Thus, there are approximately 77,047 fish in the lake after 5 years.
3. To keep the population constant, we would want $x_{n+1} = x_n$ for all n . Specifically $x_1 = x_0$. Let h be the harvesting level we wish to find. Thus, using $x_1 = x_0$ in

$$x_1 = 1.5x_0 - h,$$

we want to solve for h .

$$\begin{aligned} x_0 &= 1.5x_0 - h \\ (1 - 1.5)x_0 &= -h \\ -0.5x_0 &= -h \\ -0.5(12,230) &= -h \\ h &= 6115 \end{aligned}$$

Thus, if the lake resource managers allowed 6115 fish each season to be caught, the fish population size will remain constant from year to year.

EXAMPLE: Stocking a Lake

Consider a lake fish population whose yearly birth rate is 0.5, and yearly death rate is 0.7. Currently, there are 12,230 fish in the lake. How many fish are needed to stock the lake each year, so that the population remains constant?

Solution: The birth and death rates correspond to $\beta = 0.5$ and $\delta = 0.7$, and thus $a = 1 + \beta - \delta = 1 + 0.5 - 0.7 = 0.8$. Thus, the population is naturally decreasing. We can construct the difference equation

$$x_{n+1} = 0.8x_n + \alpha$$

where x_n is the size of the fish population at the end of year n , and α is the number of fish that are added to the lake each year. Notice that $x_0 = 12,230$. To find the number of fish needed to stock the lake each year to keep the population constant, we use $x_1 = x_0$ and solve $x_1 = 0.8x_0 + \alpha$ for α .

$$x_0 = 0.8x_0 + \alpha$$

$$\begin{aligned}
 (1 - 0.8)x_0 &= \alpha \\
 0.2x_0 &= \alpha \\
 \alpha &= 0.2(12,230) = 2446
 \end{aligned}$$

Thus, each year the lake must be stocked with an additional 2,446 fish to maintain a constant population size.

Data: Log-Log & Semi-Log Graphs

Data can be displayed with different types of scales on the horizontal and vertical axes. Frequently in plots of biological data, the horizontal axis may be labeled

$$\ln(x)$$

and the vertical axis may be labeled

$$\ln(y),$$

where x and y are the biological variables under consideration. Such a graph is called a log-log graph. If the horizontal axis is labeled

$$x$$

and the vertical axis may be labeled

$$\ln(y),$$

where x and y are the biological variables under consideration, then the graph is called semi-log. Examples of semi-log graphs can be found in the next two examples.

We will see how the log-log and semi-log graphs relate to allometric and exponential functions. We will see that if our data is close to an allometric (power) or exponential function, we can rescale the our data and see that the transformed data will look like a line on a log-log or semi-log graph.

Next, consider the exponential function $f(x) = ac^x$ where a and c are constants. Let us take the natural log of both sides of the equation and see what happens.

$$\begin{aligned}
 \ln(f(x)) &= \ln(ac^x) \\
 &= \ln a + \ln c^x \\
 &= \ln a + x \ln c
 \end{aligned}$$

Does this look like the equation of a line? Look closer using $y = \ln(f(x))$,

$$\underbrace{\ln y}_{\text{variable}} = \underbrace{\ln a}_{\text{constant}} + \underbrace{x}_{\text{variable}} \underbrace{\ln c}_{\text{constant}} .$$

We see that the y -intercept is $\ln a$ and the slope is $\ln c$ (if $c = e \approx 2.71828\dots$ then $\ln c = 1$). Notice the original vertical axis variable $f(x)$ has been rescaled to $\ln(f(x))$. Notice also that we could have just as easily used \log instead of \ln .

Now, consider the allometric (power) function $g(x) = ax^c$ where a and c are constants. Again, let us take the natural log of both sides of the equation and see what happens.

$$\begin{aligned}\ln(g(x)) &= \ln(ax^c) \\ &= \ln a + \ln(x^c) \\ &= \ln a + c \ln x\end{aligned}$$

Again, if we take a closer look we should see the equation of a line using $y = (g(x))$,

$$\underbrace{\ln y}_{\text{variable}} = \overbrace{\ln a}^{\text{constant}} + \overbrace{c}^{\text{constant}} \underbrace{\ln x}_{\text{variable}}$$

where the y -intercept is $\ln a$ and the slope is c . Here, in the allometric case, the vertical axis variable $g(x)$ has been rescaled to $\ln(g(x))$ and the horizontal axis variable has been rescaled to $\ln x$. Notice this is different from the exponential case in which only the vertical axis was rescaled.

What is the big picture?

- ★ Given an exponential equation, if we rescale the y -axis to be logarithmic, we get an equation for a line.
- ★ Given an allometric (power) equation, if we rescale both the x and y axes to be logarithmic, we get an equation for a line.

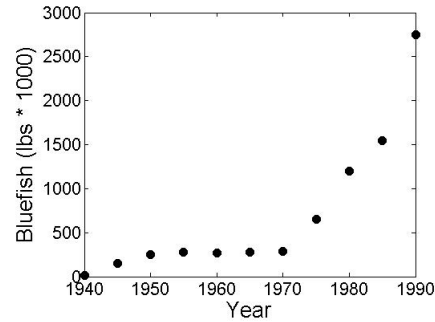
Linking Rescaling with Linear Regressions

Suppose you were given a set of data that appeared to have an exponential or allometric trend. How would you determine the exponential or allometric function that best describes that set of data? If the data had a linear trend then we could use linear regression to find the best fit line for the data. We have just seen how we can transform exponential and allometric equations to linear equations with rescaling. Using this same idea, we can rescale data so that it appears to be linear so that we may then use the techniques of linear regression to find the equation that best describes the data.

EXAMPLE:Harvesting Bluefish

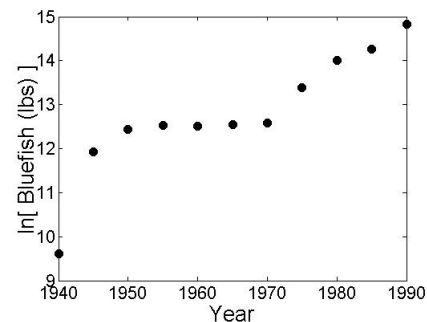
For the first example, we consider some data on Bluefish harvesting in the Chesapeake Bay, reported by the *Daily Press*, a newspaper in Virginia.

Year	Bluefish (lbs)
1940	15,000
1945	150,000
1950	250,000
1955	275,000
1960	270,000
1965	280,000
1970	290,000
1975	650,000
1980	1,200,000
1985	1,500,000
1990	2,750,000



Hopefully, looking at the data plotted out, the first observation you make is that the data are not linearly related. Let us try rescaling the y -axis data on a logarithmic scale. Additionally, on the horizontal axis, let $x = \frac{1}{5}(\text{year} - 1940)$.

x	y	$\ln y$
Year	Bluefish (lbs)	Rescaled
0	15,000	9.616
1	150,000	11.918
2	250,000	12.429
3	275,000	12.525
4	270,000	12.506
5	280,000	12.543
6	290,000	12.578
7	650,000	13.385
8	1,200,000	13.998
9	1,500,000	14.221
10	2,750,000	14.827



With the exception of the data point corresponding to the year 1940, the data appear to be more linearly related when the y -axis data is rescaled. If we fit a least-squares regression line to this data, we find the equation for the best fit line is

$$\ln y = 0.3797x + 10.8784$$

where x is the year and y is the pounds of bluefish harvested. We can convert this equation to an exponential equation:

$$\begin{aligned} \ln y &= 0.3797x + 10.8784 \\ e^{\ln y} &= e^{0.3797x + 10.8784} \\ y &= e^{0.3797x} \times \underbrace{e^{10.8784}}_{\text{constant}} \end{aligned}$$

$$y = (53,019)e^{0.3797x}.$$

Using this equation we could extrapolate how many fish were caught in 2000. First, we compute

$$x = \frac{1}{5}(2000 - 1940) = 12.$$

Then,

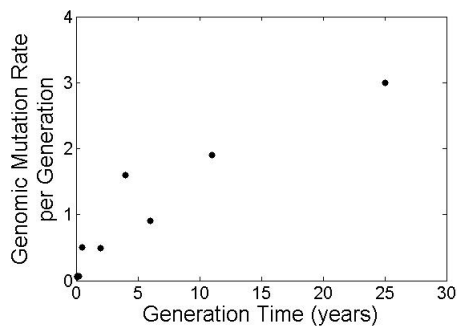
$$y = (53,019)e^{0.3797(12)} = 5,049,529.$$

Thus, we estimate that in the year 2000, about 5.05 million pounds of blue fish were caught.

EXAMPLE: Mutation Rates

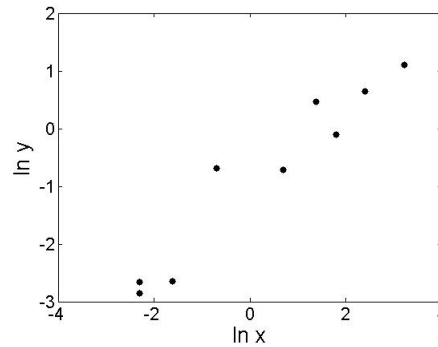
Researchers studying the relationship between the generation time of a species and the mutation rate for genes that cause deleterious effects gathered the following data.

Species	Generation Time (in years)	Genomic Mutation Rate (per generation)
<i>D. melanogaster/D. pseudoobscura</i>	0.1	0.070
<i>D. melanogaster/D. simulans</i>	0.1	0.058
<i>D. picticornis/D. silvestris</i>	0.2	0.071
Mouse/rat	0.5	0.50
Chicken/old world quail	2	0.49
Dog/cat	4	1.6
Sheep/cow	6	0.90
Macaque/New World Monkey	11	1.9
Human/chimpanzee	25	3.0



Let x = the generation time and y = the genomic mutation rate. If we rescale both the vertical and the horizontal axes then we have the following data

$\ln x$	$\ln y$
-2.3	-2.659
-2.3	-2.847
-1.6	-2.645
-0.7	-0.693
0.7	-0.713
1.4	0.470
1.8	-0.105
2.4	0.642
3.2	1.099



If we fit a least-squares regression line to this data, we find the equation for the best fit line is

$$\ln y = 0.7097 \ln x - 1.0316.$$

We can convert this equation to an allometric (power function) equation using the properties of logarithms and exponentials,

$$\begin{aligned} \ln y &= \ln x^{0.7097} - 1.0316 \\ e^{\ln y} &= e^{\ln x^{0.7097} - 1.0316} \\ e^{\ln y} &= e^{\ln x^{0.7097}} e^{-1.0316} \\ y &= x^{0.7097} e^{-1.0316} \\ y &= 0.3564x^{0.7097}. \end{aligned}$$

Using this equation, if we knew a certain species had a generation time of 5 years, we could interpolate the genomic mutation rate of this species,

$$y = 0.3564(10)^{0.7097} = 0.1827.$$

Thus, we would estimate that this particular species has a genomic mutation rate of 0.1827 mutations per generation.

Exercise 1.

Warm blooded animals use large quantities of energy to maintain body temperature because of heat loss through body surface. In fact, biologists believe that the primary energy drain on a resting warm blooded animal is the maintenance of body temperature. The table below shows the body weights in grams (g) and pulse rates in beats per minute (bpm) for nine different bird species.

Bird	Body Weight (g), x	Pulse Rate (bpm), y
Canary	20	1000
Pigeon	300	185
Crow	341	378
Buzzard	658	300
Duck	1100	190
Hen	2000	312
Goose	2300	240
Turkey	8750	193
Ostrich	71,000	65

- (a) Make a hypothesis about how you think body weight and pulse rate are related in birds? How might your hypothesis explain birds maintaining their body heat.
- (b) What type of function (linear, exponential, or allometric (power)) best describes the relationship between the weight and pulse rate data? Explain. Find the equation for the the function that best describes the data as y in terms of x .

Exercise 2.

Researchers measured the diameters of 20 trees in a central Amazon rain forest and used ^{14}C -dating to determine the ages of these trees. The data are given in the following table.

Diameter (cm), x	Age (yr), y	Diameter (cm), x	Age (yr), y
180	1372	115	512
120	1167	140	512
100	895	180	455
225	842	112	352
140	772	100	352
142	657	118	249
139	582	82	249
150	562	130	227
110	562	97	227
150	552	110	172

Consider the use of diameter x , as a predictor of age y .

- (a) Make a scatter plot of age on the vertical axis and diameter on the horizontal axis/
- (b) What type of function (linear, exponential, or allometric) best describes the relationship between the age and diameter data? Explain. Write the function for y in terms of x .

Exercise 3.

In an attempt to measure how the pace of city life is related to the size of the city, two researchers measured the mean speed of pedestrians in 15 cities by measuring the mean time it took them to walk 50 feet.

City	Population (x)	Speed (ft/s) (y)
Brno, Czechoslovakia	341,948	4.81
Prague, Czechoslovakia	1,092,759	5.88
Corte, France	5,491	3.31
Bastia, France	49,375	4.90
Munich, Germany	1,340,000	5.62
Psychro, Crete	365	2.67
Itea, Greece	2,500	2.27
Iráklion, Greece	78,200	3.85
Athens, Greece	867,023	5.21
Safed, Israel	14,000	3.70
Dimona, Israel	23,700	3.27
Netanya, Israel	70,700	4.31
Jerusalem, Israel	304,500	4.42
New Haven, CT, USA	138,000	4.39
Brooklyn, NY, USA	2,602,000	5.05

- (a) Plot the original pairs of numbers (x, y) . Is the pattern linear or nonlinear?
- (b) Compute R^2 , for the data.
- (c) Plot y against $\ln x$. Are the data more linear now than in part (a)?
- (d) Compute R^2 for y against $\ln x$. Is R^2 closer to 1 than in part (b)? What does this say about how the data are related?
- (e) Find the equation for the least squares regression line of y against $\ln x$.

Partial answers

Exercise 1

(a) As weight increases, then pulse decreases. Pulse rates can affect the maintaining of body heat.

(b) allometric, $\ln(y) = -0.28 \ln x + 7.5$ and $y = x^{0.28}e^{7.5}$

Exercise 2

allometric, $y = 1.24x^{1.23}$.

Exercise 3

- (a) Nonlinear,
- (b) $R^2 = 0.584$,
- (c) Yes, the data looks more linear,
- (d) $R^2 = 0.845$ which is close to 1 which means the original data are not linearly related,
- (e) $y = 0.410 \ln x - 0.307$.