Modeling The World – Class Notes and Exercises

John B. Little

Department of Mathematics and Computer Science, College of the Holy Cross, Worcester, Massachusetts 01610
E-mail address: `jlittle@holycross.edu`

2010 *Mathematics Subject Classification.* Primary 00-01, 92-01

Key words and phrases. mathematical modeling, linear, exponential, power law, difference equation, statistics

Abstract. These notes are the textbook – i.e. the class lecture notes and exercises – for the Montserrat Seminars MONT 100N, *Modeling the Environment*, in the Fall 2017 semester and MONT 101N, *Analyzing Environmental Data* in the Spring 2018 semester. This is a growing and developing document; you should expect modifications as the year proceeds. I will announce any changes requiring a download of a new version in class and on the course homepage. *I recommend that you keep the updated current version in electronic form on your personal computer if you are comfortable reading and referring to it that way. Unless it is absolutely necessary, please save trees and do not print out hard copies of the whole document each time a change is posted!*

Contents

Preface

As we near the end of the second decade of the 21st century, humans are grappling with a number of tough decisions concerning our place in the natural world and the consequences of how we have used various resources and impacted our environment. For instance, human populations have grown in many areas to the point that the habitats of many wild species are being lost and those species are being driven to extinction. We have continued to burn fossil fuels to generate much of the energy for our industries and transportation, with inevitable effects from the by-products and pollution that burning generates. In particular, it is growing increasingly clear that this human activity is causing long-term changes in the Earth's atmosphere and climate. New diseases have entered human populations and spread as humans have penetrated previously unsettled areas and traveled more freely around the Earth. Can we coexist with the other life forms that have evolved on this planet? Are there realistic alternatives to fossil fuels that would sustainably provide for human society's energy needs and have fewer harmful effects? How do we decide what alternatives make more sense? How do we deal with threats such as emergent diseases?

As a mathematically-literate and thinking human being, I firmly believe that our ability to develop answers to such questions and to understand the political, economic and social issues involved depends on being able to deal in an informed way with *quantitative information.* Mathematical models–equations of various sorts capturing relationships between variables involved in a complex situation–are fundamental for understanding the potential consequences of choices we make.

In the mathematical component of these courses covered in these notes, we will introduce a number of basic techniques for constructing models and see ways they can be applied to study our place in the natural world and environmental issues. More specifically, we will study the following topics in the first semester of the course:

(1) Basic techniques of measurement, data analysis, and presentation of data in numerical and graphical forms
(2) Functions and modeling – we will see how to use linear, exponential, and power functions to describe different situations and how to select an appropriate model for a given situation
(3) Difference equations and modeling – we will see how to set up and solve difference equations that describe how systems evolve over time (treating time in discrete units).

In the second semester, we will study

(1) the use of various sorts of descriptive statistics to understand the patterns in data,

(2) the basics of probability
(3) the statistics of sampling from a population, and
(4) hypothesis testing – the way scientists use statistics to demonstrate that their data supports a conclusion or interpretation.

In the main text, we will not make use of any calculus or mathematics more advanced than ideas about functions, graphs, algebra, some geometry, etc.[1] So everyone should have seen all the mathematical prerequisites and some of the basic ideas behind material we study may be familiar from mathematics courses you took in high school. What will probably be different, though, is the consistently applied and frequently environmental focus of everything we do.

However, as you surely know, understanding the predictions of even the most accurate and realistic mathematical models is *not the whole story*, not by a long shot! Some of the questions raised above have been at or near the center of the political polarization in the U.S. that has become more and more pronounced over the last 20 years.

Andrew Revkin made the following insightful and prescient observations in a 2011 New York Times blog post:

> There is not one climate dispute. There are two, and the solutions are not the same. First, we need to separate the two. The science debate does not work in politics. If you study the conservative approach to climate change policy long enough, the implication that they are trying to participate in a scientific conversation starts to fade away[2] and you realize the underlying logic they are using actually starts from the conclusion that regulation and government intervention are bad[3] ... . This allows them to make big, bold, statements about their identity and character and values rather than wallowing around in overly-precise, overly-pedantic language and data.
>
> The center-left in the U.S. has a persistent problem with this dynamic because they see every situation where they have a factual advantage as proof of their superiority and then they proceed to hammer people with logic while ignoring the repeated lessons of political strategy. The debate needs to start with values! Science has no values. Science only describes the physical world.
>
> To win the scientific debate about climate change, we just ... oh wait, we already did. But to win the political debate, we need to spend less time on the details of the scientific debate and much more on the underlying values–the costs to humanity, society, and the economy of extreme weather, local floods, local droughts, freshwater scarcity, infectious disease, food security, coastline loss, biodiversity loss, etc., etc. It sounds backwards since the political

[1] "Technical notes" sprinkled through the text and a few exercises clearly marked will deal with connections with more advanced mathematics for readers who have more background.

[2] This is true even though they phrase their arguments as attempts to sow doubt in the minds of ordinary people about the overwhelming consensus that scientists have achieved. They do this using scientific language, "alternative facts," and competing "experts" to muddy the waters.

[3] For example, this was essentially President Trump's argument for pulling the U.S. out of the Paris climate accords in June 2017 because he claimed that the accords would lead to a loss of jobs in the U.S. economy and a loss of independence for U.S. businesses in their decision-making.

> challengers are denying the possibility of those dangers, one might think we need to respond to their challenge.
>
> We do not. That's what science is for.

Check your views on science and policy, then ask yourself: are you a communitarian or an individualist? Are you a liberal, a conservative, or perhaps a libertarian? Do you believe that climate change is a "hoax" or is it a real problem that we must address in order to ensure a decent standard of life for our descendants?

While I have my own views about these questions (and you will no doubt see what they are as we discuss them), *I will never try to impose those views on you as students, or punish you, by means of grades, or comments in class, or in any other way, for holding opinions that go against mine.* The goal of this class and these notes will be to help you examine your views, whatever they are, and see if they hold up to that scrutiny and the light mathematics can shed on the facts. I won't tell you what to think, but I will expect you to give an account of your thinking and defend it with eloquence, using evidence.

John B. Little

I. Basic Quantitative Concepts

CHAPTER 1

Scales of Measurement

1.1. Introduction

Dealing with real-world data in a quantitative way means we must select one of the scientific systems of *units* for our measurements. We will begin by reviewing the *metric system*, different forms of which are used most commonly in science. We will also recall the techniques needed to convert measurements given in metric units to the perphaps more familiar English units used in the U.S. To prepare for material to come in later chapters, we will recall some of the mathematics of *logarithms* and discuss measurements on logarithmic scales.[1] Each chapter of these notes will end with an extended Chapter Project. The first of these will apply the material developed to estimate the total amount of water contained in the Greenland ice sheet and the expected increase in average sea level if all the freshwater ice sheets melt and that water is introduced into the oceans.

1.2. Metric Units

The metric system commonly used in science today was originally standardized by the First French Republic in 1799 as part of the "new start" following the revolution that deposed the king Louis XVI and abolished the French monarchy. In its basic form, the basic units were:

- The *meter*, abbreviated as m, a idealized measure of length equal to one ten millionth of the distance from the North Pole to the Equator along the line of longitude through Paris.
- Areas can be measured in square length units (e.g. square kilometers). Another unit, the *hectare*, abbreviation ha, is often used for plots of land and other everyday uses. One hectare is the area of a square 100 meters on a side, or $10,000$ square meters.
- The *kilogram*, abbreviation kg, an idealized measure of the mass of a cubical volume of pure water 1/10 of a meter on a side, at the melting point of ice.
- The *liter*, abbreviation l, is the volume given by the cube 1/10 of a meter on a side.
- The *second*, abbreviation *sec*, is the standard unit of time.
- The degree Celsius (centigrade) is the standard unit of temperature, where the freezing point of water is $0°$ C and the boiling point of water is $100°$ C.

[1] Depending on how comfortable the class is with this material, we might omit covering some or all of this in class.

To facilitate calibration, these notional definitions of the meter and the kilogram were replaced almost immediately by standard metal prototypes kept in the French national archives. But such physical models change with temperature, air pressure, and other variables. Later, in 1960, the definition of the meter was changed to make it even more precise and reproducible – a meter is now taken to be a certain number of wavelengths of a precisely specified color of light (electromagnetic radiation) that can be produced by lasers of a certain type.

For scientific work, the great advantage of the metric system is the way it follows the standard way we represent numbers – the Hindu-Arabic numerals with base-10 arithmetic. A standard series of prefixes represent positive and negative powers of 10:

Table 1. Metric Prefixes

prefix	abbrev.	power
tera	T	10^{12}
giga	G	10^{9}
mega	M	10^{6}
kilo	k	10^{3}
hecto	h	10^{2}
deca	da	10^{1}
deci	d	10^{-1}
centi	c	10^{-2}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}

As you might guess, the "hect-" in the name *hectare* comes from the multiple for 100. A hectare is 100 *ares*; the *are* is rarely used, though. Note how this system makes it possible to talk about lengths, masses, volumes, etc. over a huge range of sizes. This means that the same basic system of units applies to study both extremely small things, such as the cells in our bodies, atoms within our cells, etc. and extremely large things, such as our solar system, the Milky Way galaxy, etc. In fact, there are standard names for even larger and even smaller powers of 10 as well, but we will not have any need for them.

Example 1.1. Here are some examples of working with various metric units.

(1) a *kilogram* equals $10^3 = 1,000$ grams (a gram is roughly the mass of the plastic cap of a ballpoint pen; body masses of humans typically run in the range 50 - 100kg.)

(2) a *microgram* equals $10^{-6} = \frac{1}{1{,}000{,}000} = .000001$ of a gram (one one-millionth of a gram – roughly the mass of a large human cell like the egg cells produced in the female reproductive system)
(3) a *milliliter* equals $10^{-3} = .001$ of a liter (note or recall that a milliliter is also the same as a *cubic centimeter*, since the cube 1/10 of a meter (a decimeter!) on a side that gives the liter has a side equal to 10 centimeters, so 1 liter equals $10^3 = 1000$ cubic centimeters)
(4) a *terameter* is $10^{12} = 1,000,000,000,000$ meters (roughly the distance light travels in one hour in vacuum; the distance from the Sun to the planet Saturn is about 1.4 terameters)
(5) a *megasecond* is $10^6 = 1,000,000$ seconds (a bit over 11.5 days)

Measurements of time are rarely expressed in purely metric terms when they are communicated to humans(!) $\triangle$

If you have taken physics, you may recall that in discussions of that subject a big deal is often made of the difference between the "cgs" (= centimeters, grams, seconds) and "mks" (= meters, kilograms, seconds) versions of the metric system and what are the associated units of velocity, acceleration, work, energy, etc. depending on whether the cgs or mks system is adopted. We will not need to stress this distinction, though, and we will generally use units chosen for convenience according to the sizes of the quantities being measured.

1.3. Conversions

Because of the "powers of 10" prefix system from Table 1, conversions *within* the metric system essentially just involve shifting decimal points in numbers. This is the main reason why the metric system is so easy and why it has been adopted virtually universally both for scientific use and for everyday measurements!

Example 1.2. Here are some examples of conversions within the metric system.

(1) A distance of 17.3 kilometers can also be expressed in terms of meters like this:
$$17.3 \text{ km} \times 1000 \text{ m/km} = 17,300 \text{ m}.$$
As usual in "dimensional analysis," the units of km in the first number cancels the km in the denominator of the conversion factor 1000 meters per kilometer.
(2) Similarly, to convert a volume of 343.2 milliliters to the equivalent number of liters, we just need to remember that a milliliter is $10^{-3} = .001$ liters, and we have
$$343.2 \text{ ml} \times .001 \text{ l/ml} = .3432 \text{ l}. \qquad \triangle$$

As of 2017, there are *only three* countries in the world that do not use the metric system for everyday measurements – the African country of Liberia, the Asian country of Myanmar (formerly known as Burma), and the United States.[2] This means that the metric units above are probably less intuitively familiar than the English units:

[2]Editorial comment: this is an example of the dark side of the idea of "American exceptionalism." It can be argued that we do some things differently and better than many other nations; using the English system of units is not one of those things!

- lengths in inches, feet, yards, miles
- masses or weights in ounces, pounds, tons
- volumes in fluid ounces, quarts, gallons
- temperatures in degrees Fahrenheit

There are others too, of course! These are just the most basic and common ones. Using these units, we first need to remember conversions within the system. For instance, rounding to three significant digits in all cases we have the following.

DEFINITION 1.3. The following constants give conversions *within the English system.*

(1) 1 foot = 12 inches, so 1 inch = $1/12 \doteq .0833$ foot.[3] 1 yard = 3 feet = 36 inches. Then 1 mile = 5280 feet, so 1 foot = $1/5280 \doteq .000189$ mile.
(2) 1 pound = 16 oz., so 1 oz. = $1/16 = .0625$ pound. Then 1 ton = 2000 pounds, so 1 pound = $1/2000 = .0005$ ton.
(3) 1 quart = 32 fl. oz., so 1 fl. oz. = $1/32 \doteq .0313$ quart. Then 1 gallon = 4 quarts, so 1 quart = .25 gallon.
(4) 1 acre = $1/640 \doteq .0015625$ square mile

Then, if we want to convert between the English system and the metric system, we need to remember another set of conversion factors.

DEFINITION 1.4. The following constants give metric to English and English to metric conversions.

(1) 1 inch $\doteq$ 2.54 centimeters, and 1 cm = $1/2.54 \doteq .394$ inch.
(2) 1 foot $\doteq 12 \times 2.54$ cm $\doteq 30.5$ cm = .305 m.
(3) 1 mile $\doteq 5280 \times .305$ m $\doteq 1610$ m $\doteq 1.61$ km. Hence 1 km $\doteq 1/1.61 \doteq .621$ mile.
(4) 1 ha $\doteq$ 2.471 acres $\doteq$ .00386 square mile.
(5) 1 pound $\doteq$ .454 kilogram, and 1 kg $\doteq 1/.454 \doteq 2.20$ pounds.
(6) 1 quart $\doteq$ .946 liter, and 1 liter $\doteq 1/.946 \doteq 1.06$ quarts.
(7) To convert back and forth between a temperature F in degrees Fahrenheit and the equivalent Celsius temperature C, use

$$C = \frac{5}{9}(F - 32) \qquad \text{and} \qquad F = \frac{9}{5}C + 32.$$

Let's practice using the conversion factors from Definitions 1.3 and 1.4 on several examples.

EXAMPLE 1.5. Be sure you understand the thinking behind these and check the arithmetic as you are reading. Reading a mathematics book should be an active endeavor!

(1) To begin, let's ask how long a distance a distance of 3 kilometers is in miles, yards, and feet. From the above, we see

$$3 \text{ km} \times \frac{1}{1.61} \text{ miles/km} \doteq 1.86 \text{ miles}$$

[3]We will use the symbol $\doteq$ consistently to mean that the two quantities are *approximately equal*. The exact value of the rational number $\frac{1}{12}$ is the infinite repeating decimal $.08\overline{3}$. The approximate value .0833 is thus slightly smaller than $\frac{1}{12}$.

Then to get the equivalent length in feet, we multiply by the conversion factor from miles to feet:

$$1.86 \text{ miles} \times 5280 \text{ feet/mile} \doteq 9820 \text{ feet}.$$

Finally, to get the equivalent distance in yards, we multiply by the conversion factor from feet to yards:

$$9280 \text{ feet} \times \frac{1}{3} \text{ yards/foot} \doteq 3090 \text{ yards}.$$

(2) Next, we ask: How much does a 1 meter by 1 meter by 1 meter cube of water weigh in tons and what is its volume in fluid ounces? There are many ways to answer the first part of the question. Probably the most direct, though, is to find the mass of the water in kilograms first, and then convert this to the equivalent weight in pounds and then tons. The reason for this approach is the fact that a cube with side 1/10 meter has a mass of 1 kilogram by the definition of the metric units. Since our cube has a side that is 10 times as long as this, the mass of the water is $10 \times 10 \times 10 = 1000$ kilograms. Then

$$1000 \text{ kilograms} \times 2.20 \text{ pounds/kilogram} \times \frac{1}{2000} \text{ tons/pound} \doteq 1.1 \text{ tons}.$$

This says, for instance, that just the water contained in a 3 meter by 2 meter by 1/6 meter waterbed would also weigh about 1.1 tons![4] For the second part, we note that the volume is $10 \times 10 \times 10 = 1000$ liters, and

$$1000 \text{ liters} \times 1.06 \text{ quarts/liter} \times 32 \text{ fl.oz./quart} \doteq 33920 \text{ fl.oz.}$$

(3) Now, suppose we have a flat 1 square mile field that is flooded with water to a depth of one inch. We ask: what is the total volume of the water in cubic meters? Thinking of the water as (approximately!) a rectangular solid, the volume is length $\times$ width $\times$ height. We want the volume in cubic meters, so it makes sense to convert all three dimensions to meters first, then compute the volume in cubic meters. So we have 1 mile $\doteq$ 1610 meters, and

$$1 \text{ inch} \times .083 \text{ foot/inch} \times .305 \text{ meter/foot} \doteq .0235 \text{ meter}.$$

Hence the total volume of the water is approximately

$$1610 \text{ meters} \times 1610 \text{ meters} \times .0235 \text{ meter} \doteq 65580 \text{ cubic meters}.$$

We treated the shape of the field as though it was a perfectly flat square one mile on a side. The given information says we may ignore the curvature of the Earth here and assume the field is perfectly flat (that is, contained in one plane). This is reasonable because one mile is so much smaller than the Earth's radius (approximately 4000 miles). The answer would also be the same no matter what the actual shape of the field was because a solid of uniform height over a fixed perfectly planar base always has volume equal to the area of the base times the height.

[4]Think about that before setting one up in your 4th-floor apartment bedroom!

(4) Finally, a temperature of 35° Celsius is equivalent to a Fahrenheit temperature of

$$\frac{9}{5} \cdot 35 + 32 = 95^\circ \text{ F}.$$

The (much more comfortable) Fahrenheit temperature 72° F is equivalent to a Celsius temperature

$$\frac{5}{9}(72 - 32) \doteq 22.2^\circ \text{ C}. \qquad \triangle$$

1.4. Estimation

In many circumstances, exact measurements of quantities in real-world processes may be unavailable due to the difficulty or the cost of carrying them out. In these situations, a good *estimate* or "educated guess" (with emphasis on the "educated," of course!) may be the most we can hope for. Making good estimates requires both a solid understanding of the process and an intuitive grasp of the units involved. Here is a simple example.

Example 1.6. Suppose we wish to estimate the average per capita daily usage of water for a U.S. resident. First we must think of all the ways that we use water in a typical day. Most of us probably use

- about 16 gallons to take a shower lasting 8 minutes (at about 2 gallons per minute)
- about 24 gallons in bathrooms (halve that if you use efficient 1 gallon per flush toilets, but most people *do not* have them)
- about 2 gallons for drinking and cooking purposes

That's about 42 gallons a day, but are those the only uses of water most of us have? If you think about it, you'll see that there are large water uses in addition to these:

- if you water a lawn or use a swimming pool, add 25 gallons a day
- add 4 gallons per day for use in washing laundry
- add 4 gallons per day if you use a dishwasher

This adds up to an estimate of about 75 gallons per day (and it doesn't even include uses of water to wash the car or clean the house, etc.)

In fact, this is pretty close to the mark. The U.S. Environmental Protection Agency estimates that the average American household uses about 300 gallons of water per day[5].

By way of contrast, per capita water usage in most other parts of the world (even in economically advanced areas such as Western Europe) is significantly lower. Even though most parts of the U.S. have adequate water resources to support our lifestyles at the present time, the same is not true in drier parts of the world. The cost of purifying water for human consumption is also increasing. $\triangle$

1.5. A Feature of Human Perception

To introduce our next topic, we will discuss an interesting feature of the way our senses (vision, hearing, taste, touch, smell) deal with stimuli from the physical

[5]According to `https://www.epa.gov/watersense/how-we-use-water`

FIGURE 1.1. Circles with Areas in Arithmetic Progression

world. We will illustrate our point with two series of images of circles. In Figure 1.1 above, the areas of the circles (in suitable square units!) would be numbers in the *arithmetic progression* $1, 2, 3, 4, 5, 6, 7, 8$. What this means is that the change in area from each circle to the next is always the same: 1 square unit.

Now compare this with the sequence of circles in Figure 1.2 on the following page. Here, if the smallest circle at the left had area 1, then areas would be $1, 2, 4, 8, 16, 32, 64, 128$. (The scale is not the same as in the first figure because the range of areas is much greater.) In other words, each circle is twice as large in area as the one before it. We call such a sequence of number values a *geometric progression*.

Compare the two sequences of circles carefully. In the first case (Figure 1.1), you should notice that the *rate of growth* in the areas as you sweep your vision from the left to the right seems to get less and less. In fact, there is apparently very little difference between the last two circles: if presented with those two circles in isolation you might be hard pressed to see the difference at all (even though their areas differ by the same amount as the areas of the first two circles). On the other hand, the growth of the areas in Figure 1.2 is seemingly steadier and we have no difficulty in perceiving that the areas are increasing.

Psychologists call the phenomenon we are seeing here the *Weber-Fechner law* of perception. In rough terms, the Weber-Fechner law says that human sense perception works on a *logarithmic* scale. We will explain what this means in detail in the next sections.

1.6. LOGARITHMS

You have probably seen logarithms in your high school algebra or precalculus classes. Recall that the idea is the following. Given a positive number $a \neq 1$, called the base of the logarithms, and any positive number x we say that

$$\boxed{y = \log_a(x) \quad \text{if (and only if)} \quad a^y = x.} \tag{1.1}$$

In other words, the base a logarithm of x is the *exponent* to which a must be raised to yield the number x.[6]

[6]The reason for the restriction $a \neq 1$ can be seen if we think of trying to find y with $1^y = x$ if $x \neq 1$. In practice, values $a > 1$ are used much more commonly than values $0 < a < 1$.

FIGURE 1.2. Circles with Areas in Geometric Progression

EXAMPLE 1.7. For instance

$$\log_3(81) = 4 \quad \text{since} \quad 3^4 = 81.$$

Similarly,

$$\log_{10}\left(\frac{1}{1000}\right) = \log_{10}(10^{-3}) = -3.$$

For numbers x that are not exactly equal to whole number powers of the base a, a calculator, mathematical computer software, or a table is used to find values of the logarithm $\log_a(x)$. For example by any of these methods, we find

$$\log_{10}(2) \doteq .30103.$$

This is true since

$$10^{.30103} \doteq 2.$$

Similarly, using a calculator

$$\log_{10}(.58) \doteq -0.23657.$$

Because $.58 < 1$, a negative exponent y is needed to produce an equation $10^y = .58$. $\triangle$

From (1.1), if we know $y = \log_a(x)$, then recovering x is simply a matter of computing $x = a^y$. For instance, if we know $\log_{10}(x) = 4$, then $x = 10^4 = 10,000$.

Properties of exponents carry over into corresponding properties for logarithms.

PROPOSITION 1.8. *For all $a > 0$ different from 1 and all positive x, x_1, x_2, b, the base-a logarithms satisfy*

(1) $\log_a(x_1x_2) = \log_a(x_1) + \log_a(x_2)$ – *in words, "the log of a product is the sum of the logs."*
(2) $\log_a\left(\frac{x_1}{x_2}\right) = \log_a(x_1) - \log_a(x_2)$ – *in words, "the log of a quotient is the difference of the logs."*
(3) $\log_a(1) = 0$.
(4) $\log_a(x^b) = b \cdot \log_a(x)$ – *in words, "the log of a base to a power is the power times the log of the base."*

PROOF. We are not going to be doing a lot of proofs, but we will look at these because understanding how logarithms work is important and the proofs show that. In fact, all of these statements follow from properties of exponents.

(1) For instance, if $y_1 = \log_a(x_1)$ and $y_2 = \log_a(x_2)$, then we have equations

$$a^{y_1} = x_1 \quad \text{and} \quad a^{y_2} = x_2.$$

Hence multiplying and using the fact that the exponents add when we multiply two powers of the same base a, we have

$$x_1 x_2 = a^{y_1} a^{y_2} = a^{y_1+y_2}.$$

Therefore

$$\log_a(x_1 x_2) = y_1 + y_2 = \log_a(x_1) + \log_a(x_2),$$

since that is the exponent in the equation giving $x_1 x_2$ as a power of a.

(2) This follows in the same way since

$$\frac{x_1}{x_2} = \frac{a^{y_1}}{a^{y_2}} = a^{y_1-y_2}.$$

(3) The equation $\log_a(1) = 0$ follows since $a^0 = 1$.

(4) If $y = \log_a(x)$, then $a^y = x$. Hence, raising both sides to the b power, we get $x^b = (a^y)^b$. But $(a^y)^b = a^{b\cdot y}$ since the exponents multiply. This gives $\log_a(x^b) = b \cdot y = b \cdot \log_a(x)$.

□

The idea of logarithms is usually ascribed to the Scottish mathematician John Napier (1550 - 1617). Napier was interested mostly in the ways parts (1), (2) and (4) of Proposition 1.8 can be used to simplify numerical calculations. The idea is that you have a table of logarithms computed for a suitably "dense" set of numbers between 1 and 10, say including $\log_{10}(2.31) \doteq 0.3636120$. Then you can deal with numbers of any magnitude by using scientific notation like this:

$$\log_{10}(2.31 \times 10^4) = \log_{10}(2.31) + \log_{10}(10^4) = \log_{10}(2.31) + \log_{10}(10^4) \doteq 4.3636120.$$

Most importantly, using this approach, complicated jobs of multiplying, dividing, or raising numbers to powers can be replaced by the simpler computations of addition and subtraction of the logarithms, or multiplication of the logarithm by the power. As recently as the early 1970's, when I was in high school,[7] it was still possible to find whole semester courses devoted to these calculations in some school curricula. The textbooks usually contained 7-place log tables making up most of the book(!) Of course the availability of calculators has made the very idea of such courses hopelessly old-fashioned today(!)

Most scientific calculators have keys for computing both $\log_{10}$, the base-10 or *common logarithm*, and ln, the so-called *natural logarithm*, where the base is a number called $e \doteq 2.71828$. The reasons for calling this odd-looking choice of a "natural" are studied in calculus classes and we will not go into the details here.

Much scientific work uses the base-10 or common logarithms pretty exclusively and we will follow that practice in these notes as well. You should be aware, however, that converting back and forth between these two (and actually *any* two) systems of logarithms is very easy if you ever need to do it. The reason is a consequence of the reasoning in part 4 of Proposition 1.8.

[7] I hesitate to admit this, but that was before the advent of electronic calculators!

For instance, if $e^y = x$, so $y = \ln(x)$, then using the equation $e = 10^{\log_{10}(e)}$ and substituting in for the e in $e^y = x$, we also have

$$(10^{\log_{10}(e)})^y = 10^{y \log_{10}(e)} = x,$$

so

$$\log_{10}(x) = y \log_{10}(e) = \ln(x) \log_{10}(e).$$

In other words, to convert from the natural log of x to the common log of x, you just multiply by the constant $\log_{10}(e) \doteq .43429$. To convert the other way, you divide the common log by .43429, which is the same as multiplying by $2.3026 \doteq \ln(10)$.

The most general statement along these lines is given in the following general conversion formulas. For any positive numbers a, b, x:

$$\boxed{\log_b(x) = \log_a(x) \cdot \log_b(a) \quad \text{and} \quad \log_a(x) = \frac{\log_b(x)}{\log_b(a)}.} \tag{1.2}$$

Here is an example.

Example 1.9. Say $b = e$, so $\log_b = \ln$ is the natural logarithm. Also take $a = 10$. Using a calculator, we have

$$\ln(4.33) \doteq 1.4656.$$

Hence rounding to 5 decimal digits,

$$\log_{10}(4.33) \doteq \frac{1.4656}{\ln(10)} \doteq \frac{1.4656}{2.3026} \doteq .63650.$$

All the values are rounded here; if you compute $\log_{10}(4.33)$ directly you will get a value closer to .63649. These small differences generally do not make much difference when dealing with real-world data, where a measurement may only be known to within two or three significant digits anyway due to possible experimental or observational errors. But if more precise values are required, it is always possible to use all of the decimal digits returned in the calculator values of logarithms. △

1.7. Logarithmic Scales

We say a quantity or a magnitude is defined using a *logarithmic scale* if it is computed using a logarithm of some other quantity. For instance, if instead of plotting values of a distance along the x-axis in a plot we used the logarithms of the distances, then we would be using a logarithmic scale.

Example 1.10. The Weber-Fechner law we discussed before can be stated in a more precise way by saying that the human visual system deals with information about the areas of plane regions using a logarithmic scale. That is, it is effectively the logarithm of the area of the region that "registers" more directly in our minds than the area itself. Recall that in Figure 1.1, we saw circles of areas $1, 2, 3, 4, 5, 6, 7, 8$. If we take the (common) logarithms of these values, we get the numbers in Table 2. The successive differences are decreasing as we go to the right. For instance, the first two logs differ by .30103, while the last two differ by much less, namely approximately .06. Thus, the logarithmic scale seems to be reproducing the intuitive idea that the last two areas differed much less in visual terms than the first two.

TABLE 2. Logarithms of the Areas in Figure 1.1

Area	1	2	3	4	5	6	7	8
$\log_{10}$(Area)	0	.30103	.47712	.60206	.69897	.77815	.845100	.90309

On the other hand, in Figure 1.2, we saw circles of areas $1, 2, 4, 8, 16, 32, 64, 128$. The common logarithms of these numbers are (approximately) shown in Table 3. Here the successive differences are all $\log_{10}(2) \doteq .30103$ and this seems to match the perception that those circles are growing steadily in size. $\triangle$

TABLE 3. Logarithms of the Areas in Figure 1.2

Area	1	2	4	8	16	32	64	128
$\log_{10}$(Area)	0	.30103	.60206	.90309	1.20412	1.50515	1.80618	2.10720

There are many other important examples of logarithmic scales used in science. We will discuss one of these next, and you will see others in the exercises and in future chapters.

EXAMPLE 1.11. In chemistry, the *pH* (from "potential of hydrogen") of an aqueous (water-based) solution is a measure of its acidity or alkalinity. Acids and bases are distinguished by the level of H^+ ions present. Formally,

$$\boxed{pH = -\log_{10}(a_{H^+}) = \log_{10}\left(\frac{1}{a_{H^+}}\right),} \tag{1.3}$$

where a_{H^+} is the *hydrogen ion activity* (essentially the concentration in units of moles per liter of solution). For instance pure water has $a_{H^+} \doteq 1.0 \times 10^{-7}$, and hence

$$pH = -\log_{10}(1.0 \times 10^{-7}) = 7.$$

Acids are aqueous solutions with $pH < 7$ and the smaller the pH, the "stronger" the acid. For instance, lemon juice has $pH \doteq 2$, while concentrated hydrochloric acid can have $pH \doteq 0$. At the other end of the scale, on average, human blood is slightly alkaline or basic ($pH \doteq 7.4$) and strongly basic solutions like extra-strength Drano (a liquid drain cleaner) can have $pH \doteq 14$. Because of the $\log_{10}$ in the definition (and the negative sign), if solution A has $pH = 4$ and solution B has $pH = 5$, then the hydrogen ion activity for solution A is 10 times as large as that for solution B. The a_{H^+} for solution A is 1.0×10^{-4} and the a_{H^+} for solution B is 1.0×10^{-5}. $\triangle$

1.8. Chapter Project

The ice sheet covering the island of Greenland is currently one of the major concentrations of frozen fresh water situated over a land mass on the Earth. It is estimated to contain about $\frac{1}{9}$ of the total freshwater ice on the planet. Because this water is presently on land, if it should melt, that water would add to the depth of the oceans. This is different from the situation for the Arctic sea ice, or much of the ice in the shelves surrounding Antarctica, which rest on sea water. Melting in those cases would not appreciably change sea levels because that ice is floating on and supported by water already.[8] There are important freshwater ice sheets covering the land mass of Antarctica as well as these sea ice formations.

As we will see, there is a large volume of water contained in the Greenland ice sheet. For this reason, its fate is of more than passing interest for all humans–many of our major cities and settled coastal areas occur in regions that are close enough to the current sea level that any significant increase will cause major disruptions. For instance, the *highest* (natural) elevation of any point on the island of Manhattan in New York City is only 81 meters ($\doteq$ 265 feet) above sea level and most of the island lies much lower than that. The effects of the relatively small increase in average sea level that have already occurred were evident, for instance, in the flooding of Manhattan that happened during "super-storm Sandy" in October, 2012.

The news on this front is not encouraging because major melting has been observed on Greenland in recent summers. This is thought to be a result of a combination of higher air temperatures and decreases in the reflectivity of the ice due to deposition of soot particles from sources on other continents (this increases the amount of energy absorbed from solar radiation during the summer months and increases the rate of melting).

Refer to the map of Greenland showing the depth of the ice sheet covering most of its land area in Figure 1.3. We want to use the information here to estimate the volume of the water contained in this ice sheet and understand the possible (now maybe even probable!) effects if it melts completely.

Our methods will yield rough estimates or approximations of the ice volume. They will be based on the following simplifying assumptions:

- Even though it appears as a large area on the familiar Mercator projection maps you may have seen, that map projection distorts areas of regions near the poles and makes them look much larger than they actually are. Greenland does not make up a very large a portion of the surface area of the (roughly spherical) Earth. As a result we will not lose too much if we simply estimate areas as though they corresponded to areas on the flat map (i.e. without trying to take the curvature of the Earth into account). The distance scale marked in the legend of the map can be used, together with a ruler, to approximate linear dimensions of regions.
- The ice depths (in meters) of the various portions of the ice sheet are encoded in the map by the 6 different colors. For the purposes of this exercise, let's take the ice depth in a region to be *constant* at the *lower limit* of the range of depths shown in the legend of the map. For example, this will mean that all points colored with the darkest shade of blue will have ice depths equal to 3000 m (even though the actual depths can be larger). Note that this will produce estimates on the small side of the actual depth.

Questions. The chapter project will involve investigating the following questions and writing up your results as directed below.

[8]This is good since as of July 12, 2017 the Larsen C ice shelf, with an area of about 44,200 square kilometers has broken off from the Antarctic coast and become the largest iceberg ever observed by humans.

FIGURE 1.3. The Greenland Ice Sheet, as of about 2013, source https://en.wikipedia.org/wiki/Greenland_ice_sheet

(A) Estimate the areas of each of the six different ice sheet regions identified by the colors in the map. Explain how you are doing this in a clearly-written paragraph. Note: There are *many ways* to do this in a reasonable fashion and there is not a single correct answer! One suggestion would be to print out paper copies of the page with the map and draw in collections of rectangles covering, or nearly covering, each of the regions by hand. You don't need to get super-detailed or picky, but be as accurate as is reasonably possible. Then use the distance scale to estimate the areas of those rectangles and add up a total for each region.

(B) Multiply each of your area estimates by the depth estimate to get a volume estimate. Add the ice volume estimates to get a total volume and express in units of *cubic kilometers*. (As a "reality check" for your method, the total volume of the Greenland ice sheet is often estimated to be about $3,000,000$ cubic kilometers. How close did you come to that?) Recall that the Greenland ice sheet accounts for about $\frac{1}{9}$ of the total freshwater ice on Earth. What is your estimate of the total volume of freshwater ice on Earth?

(C) Now imagine that an amount of water equal to your total volume estimate is added to the oceans all at once. How much would sea levels rise as a result? One way to estimate that is to use the same idea as in part (3) of Example 1.5 above. What is the total surface area of the oceans on the Earth? (You should look this up online. If you find different estimates, how will you choose which one to use? Explain your thinking.) If the water from the melted Greenland ice sheet was spread evenly over that area, how deep would it be, in meters? Would the actual change in sea level be less than or greater than this estimated height? Explain.[9]

(D) Find an elevation contour map of Manhattan. Use that information to estimate what portions of that island would be under water if all the freshwater ice sheets melted.

Write up your solutions for these questions as a project report. Include all of your calculations, the version of the map you used to estimate the areas of the

[9]Technical note: In case you are worried about the fact that we are ignoring the spherical shape of the Earth, this method is actually sufficient (i.e. accurate enough) for our purposes because of the fact that the depth of the water would be much smaller than the radius of the Earth. Here's one way to think about it: If there is no land area, then adding the water from the melted ice sheet is equivalent in mathematical terms to changing the radius of the spherical Earth from r to $r + \Delta r$ (with Δr representing the depth of the new water, much smaller than r itself). The volume of the added water is the difference between the volume of the larger sphere of radius $r + \Delta r$ and the volume of the original sphere of radius r:

$$\frac{4\pi(r+\Delta r)^3}{3} - \frac{4\pi r^3}{3} = 4\pi r^2 \times \Delta r + 4\pi r \times (\Delta r)^2 + \frac{4\pi(\Delta r)^3}{3}.$$

Since we assume Δr is much smaller than r, then the last two terms on the right are negligible in size compared to the first term and we obtain an estimate

$$\text{volume of added water} \doteq 4\pi r^2 \times \Delta r = \text{surface area of sphere} \times \Delta r.$$

The change in sea level is then approximated by

$$\Delta r \doteq \frac{\text{volume of added water}}{\text{surface area of sphere}}.$$

The same idea works even if the water covers only a portion of the surface area of the sphere.

different regions of the ice sheet, and your answers to all the "explain" portions of the questions above.

Chapter Exercises

(1) Express a volume of 5343 cubic centimeters in terms of liters and then cubic meters.

(2) Express a weight of 4.3 tons in pounds and then in ounces.

(3) Express the volume 8 fl. oz. in terms of liters and then milliliters.

(4) Express an area of 130 square kilometers in terms of square miles, square yards, and finally square feet.

(5) On July 12, 2017, the Larsen C ice shelf broke free of the rest of the Antarctic ice shelves and became a free-floating iceberg. Its total area at the time was estimated to be 44,200 square kilometers (roughly the size of the state of Delaware!)

   (a) What was the area of the new iceberg in square miles?

   (b) The average thickness of the Larsen C ice shelf was estimated at 350 meters. Estimate the total volume of ice contained in the new iceberg at the time it separated ("calved"), in cubic kilometers, then in cubic feet.

(6) How many minutes have passed since the inauguration of Donald Trump as U.S. President at 12:00noon on January 20, 2017? (Don't estimate – find a value as close to exact as possible. Your answer will depend on when you are doing this problem, of course. State what that time is!)

(7) Light travels at a velocity $c \doteq 299,792,458$ m/sec in vacuum.

   (a) What is the equivalent velocity in units of mi/hr? (Note: to get from m/sec to mi/hr, you will need to multiply by the conversion factors for mi/m and then sec/hr.)
   (b) How long is a *light year* (the distance light travels in one year), measured in km, then in mi?
   (c) At a speed of $.5c$, how long would it take to get from Earth to Proxima Centauri, the nearest star outside our solar system? (Look up the distance online.)

(8) There is exactly one temperature where the Celsius and Fahrenheit scales give the same numerical value. What is this temperature? Show your work to determine it.

(9) Estimate how many times an average human heart beats over the course of the lifetime of its "owner." Your number should be amazing if you think about it–very few things we can make with moving parts are that durable! (Hints: First estimate how many times the heart beats over a short time span like one minute. This depends on lots of things–whether the person is at rest or exercising, what the general health of the person is, how old the person is, etc. Don't worry too much about those, though. Most people spend more time in a state close to resting than in intense exercise. You'll also need to estimate

the average life span of a human being. Feel free to look up any information you need online, but think carefully about what you find and ask whether it is general enough for your purposes!)

(10) Without using a calculator, compute exactly:

(a) $\log_5(125)$

(b) $\log_3\left(\frac{1}{729}\right)$

(c) $\log_{10}(.001) \times \log_7(49)$

(d) $\log_2\left(\frac{16^3 \times 8}{1024}\right)$. Do this using Proposition 1.8 first, then check your work by simplifying the fraction $\frac{16^3 \times 8}{1024}$ first before taking the logarithm.

(11) Using a calculator (and converting as necessary using (1.2)) find

(a) $\log_{10}(5.34689)$, then $\log_{10}(53.4689)$, then $\log_{10}(534.689)$. Explain the pattern you are seeing using Proposition 1.8. Note that the decimal point is shifting one space to the right each time.

(b) $\log_7(34.333)$

(c) $\ln(100.3)$

(12) (Refer to (1.3).) What is the hydrogen ion activity a_{H^+} for a solution with $pH = 8$? Same question if $pH = 3$?

(13) The intensities of sounds are often measured in units called *decibels*, or dB. The method involves comparing the sound pressure level measured as a result of the sound with a standard reference level corresponding to "silence," where only the resting air pressure is experienced by the measuring device. In acoustics, the following definition is used for the sound pressure level, L, measured in decibels:

$$L = 20 \times \log_{10}\left(\frac{p_m}{p_r}\right).$$

Here, p_r is the reference sound pressure $p_r = 20$ micropascals[10] in air and p_m is the measured sound pressure. From this formula, we can see that sound pressures are measured along a *logarithmic scale*, in agreement with what we said before about the Weber-Fechner law and human sense perception.

(a) What is the sound pressure level in decibels corresponding to a measured sound pressure $p_m = 5000$ micropascals?

(b) What is p_m if L has the value 1 dB, 2 dB, 10 dB?

(c) A jet engine produces sound pressure at the level of 150 dB. What would the measured sound pressure p_m be then?

(14) Strength of earthquakes is currently measured using the *moment magnitude* scale. The moment magnitude M is a quantity defined as

$$M = \frac{2}{3}\log_{10}(S) - 10.7$$

[10] The pascal is a metric system unit with dimensions of force per unit area. A pressure of 1 pascal is 1 newton per square meter. The pascal has dimensions kg/(m $\times$ sec^2).

FIGURE 1.4. Kudzu vines in Atlanta, GA

where S is the seismic moment (a quantity measured in units of force times distance and representing the energy released). This quantity is also given on a logarithmic scale. The constants $\frac{2}{3}$ and -10.7 are chosen so that the numbers generated are roughly equal to the older *Richter scale* magnitude, which used a different method.

(a) If two earthquakes have seismic moments S_1 and S_2 and moment magnitudes M_1 and M_2, show that the ratio between the two seismic moments is
$$\frac{S_1}{S_2} = 10^{\frac{3}{2}(M_1 - M_2)}.$$

(b) If one earthquake has moment magnitude $M_1 = 6$ and a second one has moment magnitude $M_2 = 5$, how much stronger is the first one in terms of the ratio between the two seismic moments? Express your ratio in decimal form and explain its meaning.

(15) (For the more mathematically minded.) Look at the argument given in footnote 9 on page 16. If Δr is much smaller than r, why can we say that $t_2 = 4\pi r \times (\Delta r)^2$ is negligible compared to $t_1 = 4\pi r^2 \times \Delta r$? Similarly, why can we say $t_3 = \dfrac{4\pi(\Delta r)^3}{3}$ is negligible compared to t_1? Hint: What are the ratios $\frac{t_2}{t_1}$ and $\frac{t_3}{t_1}$ after you simplify?

(16) *Invasive species* are a problem in many habitats. If you have spent any time in the southeastern part of the U.S. (especially South Carolina, Georgia, Alabama) you have surely seen one of the most successful and persistent invasive plant species–*kudzu* vines (*Pueraria montana*), "the vines that ate the South." See Figure 1.4.[11] Like many of the most notorious examples of invasives, kudzu was intentionally introduced by humans into its present habitat. People imported kudzu into the U.S. from Japan and other parts of Asia in the late 1800's. The Soil Erosion Service and Civilian Conservation Corps promoted its use as a ground cover to prevent soil erosion; other people sold the plants as ornamental ground cover. Unfortunately, the climate and environment of the South were a perfect match for kudzu and it has proliferated. The major problems caused by kudzu come from the fact that it is what has been called a "structural

[11]Source: `https://en.wikipedia.org/wiki/Kudzu_in_the_United_States`, downloaded June 28, 2017.

parasite"–rather than supporting itself, it covers and smothers human-built structures (including power and telephone lines), and other trees and plants. It has very rapid growth and a versatile "double-barrelled" reproductive strategy. Kudzu uses both sexual reproduction via flowers and seeds, and asexual self-cloning whereby a kudzu plant generates vines that root, separate from the parent plant, and grow into new genetically identical individuals.

(a) There were no kudzu plants in the U.S. before 1876. Today it covers about $3,000,000$ hectares (about $7,400,000$ acres). How much is that in square miles?

(b) (Preview of Chapter 2.) What is the average growth rate of the area covered by kudzu since 1876 in hectares per year, and also in acres per year?

(c) Another advantage of kudzu is that it has a very efficient symbiotic relationship with soil bacteria that lets the plants fix large quantities of nitrogen as ammonia (NH_3) in the soil it grows in. This lets kudzu plants essentially make their own fertilizer(!) and grow initially in areas that are too nitrogen poor to support other plants. According to data reported by Forseth and Innis[12] a stand of kudzu can fix 235 kg of nitrogen per hectare per year. How much is that in pounds per square mile per month? Not all of this ammonia stays in the soil either; it can leach into streams and lakes and cause changes in the plants and animals there as well.

[12] Kudzu (*Pueraria montana*): History, Physiology, and Ecology Combine to Make a Major Ecosystem Threat, Critical Reviews in Plant Sciences vol. 23 (2004) online, consulted June 28, 2017.

CHAPTER 2

Ratios, Percents, Proportions

2.1. Introduction

Measurements sometimes have meaning in isolation. For instance, on June 21, 2017 as I was writing this in Worcester, MA, the day was sunny and the high temperature was 81° F. You can see from that information that it was a very comfortable early summer day(!) However, there are also times when we want to *compare* measurements or other quantities in various ways. Here is an everyday example that will illustrate some of the key points we want to make in this chapter.[1]

According to a news story published in the *New York Daily News* on January 4, 2017[2] there were 352 homicides reported in New York City during the calendar year 2015 versus 335 homicides during the calendar year 2016. In looking at the homicides from one year to the next, the most basic comparison would be to say $335 - 352 = -17$, so the number of homicides *decreased by* 17. But we could also ask more about this situation: Is that a big decrease or a small decrease? One way to answer that more refined question would be to compute what is called the *percent decrease*. The idea here is that we look at the decrease *as a fraction of the number of homicides in the previous year*, then express that fraction as a *percent* by giving it as a number out of the standard nominal total of 100%.[3] The computation would be

$$\text{percent decrease} = \frac{335 - 352}{352} \times 100\% \doteq -4.8\%. \tag{2.1}$$

A decrease of 4.8% would usually be interpreted as a "medium size" decrease–not huge, but not negligibly small either. The homicide figures were moving in the right direction in 2016, but the change was incremental rather than massive.

On the other hand,[4] there were 181 homicides in the state of Arkansas in calendar year 2016 versus the 335 homicides reported in New York City that year. Looking at the New York City homicide figure versus the Arkansas state homicide figure, it might be tempting to think that New York City was measurably less safe and more crime-ridden than Arkansas. After all, 335 is only a bit less than twice the Arkansas figure since $2 \times 181 = 362$, and the 181 is the figure for a whole state, not just one city!

[1] Depending on how comfortable the class is with this material, we might omit covering some or all of this in class.

[2] Archived at `http://www.nydailynews.com/new-york/nyc-crime/nyc-historically-number-shootings-2016-article-1.2933098`, consulted June 21, 2017.

[3] This is the origin of the word "percent" since 100 is "centum" in Latin. The % sign was invented for use in accounting and in specification of interest rates.

[4] According to `https://www.neighborhoodscout.com/ar/crime`, consulted June 21, 2017.

But we have to ask: Is this is a reasonable way to make the comparison? Or might we be skewing the data in some way by just comparing the numbers of homicides? Before you read on, ask yourself: Might we be leaving out some relevant information?

In 2016, the population of Arkansas was approximately 2.98 million people, while the population of New York City was approximately 8.54 million people. If we take that difference into account by computing the *per capita homicide rates*, that is, the number of homicides per person[5] in the populations, we get approximately the following figures:

$$\text{New York per capita homicide rate} = \frac{335}{8.54 \times 10^6} \doteq .000039,$$

and

$$\text{Arkansas per capita homicide rate} = \frac{181}{2.98 \times 10^6} \doteq .000063.$$

These are both quite small numbers, of course, since the number of homicides was a very small fraction of the total population in either case. However, there is also a comparison we could make between the two per capita rates by computing a *percentage difference*–that is by expressing the difference between the per capita rates in Arkansas and New York City as a fraction of the per capita rate in New York City, then converting that to a percentage. By that measure, the rate in Arkansas was actually

$$\frac{.000063 - .000039}{.000039} \times 100\% = 61.5\% \tag{2.2}$$

higher than the rate in New York City. This percentage difference would be usually be interpreted as being relatively high. By this measure, New York City was actually quite a bit less dangerous than Arkansas![6]

Another possible way to interpret these numbers is that in New York City, if you picked a person at random from a list of the population, then there would be a .0039% chance that person was murdered in 2016. Similarly, there was .0063% chance a randomly chosen Arkansan was murdered. One possible objection to this way of thinking about the numbers, though, is that we don't know whether all of the homicide victims were natives or whether there might have been visitors included in those homicide totals. So this way of thinking about the numbers is slightly suspect in this particular case. Also, to make a fairer comparison between crime levels in the two locations, we would probably also want to look at rates of other violent and non-violent crimes, not just homicides. But we could use the same methods to make those comparisons too and probabilities are often computed by the same sort of ratios that we have been discussing in this example. In this chapter, we will review the ideas of ratios, proportions, and percents that we have been using here and see many of their applications.

[5] "Capita" comes from the Latin word for "head."

[6] I'm not trying to pick on Arkansas here. There are a number of other relatively small states where the comparison would work out similarly!

2.2. Ratios and Percentages

From the mathematical point of view, taking a ratio always involves *dividing* one numerical measurement by another numerical measurement. The units of the two measurements can be the same or different, depending on the situation.

If the units of the two quantities are the same, then the units in the numerator and the denominator of the ratio cancel, and the value of the quotient gives what is called a *dimensionless quantity*, or pure number. The meaning of the ratio in this case is usually just a simple comparison of the sizes of the numerator and denominator in the quotient.

Example 2.1. The formula for computing sound pressure levels

$$L = 20 \times \log_{10}\left(\frac{p_m}{p_r}\right).$$

from Exercise 13 in Chapter 1 involves the quotient $\frac{p_m}{p_r}$ where the reference pressure p_r and the measured pressure p_m have the same units (micropascals). Hence the quantity L will also be a pure number, or dimensionless quantity. △

Percentages are also frequently used to represent one quantity as a fraction or multiple of some other reference quantity, measured with the same units. Any ratio where the numerator and denominator have the same units can be converted into an equivalent percentage of the denominator value simply by multiplying the numerical value of the quotient by 100%. The effect is to convert the ratio into another equivalent ratio with an (implicit) denominator of 100. Thus if a first quantity is 25% of a second quantity, we see that the first is

$$\frac{25}{100} = \frac{1}{4}$$

of the second. Similarly, if the first quantity is 150% of the second, the first is

$$\frac{150}{100} = \frac{3}{2}$$

of the second, or "half again as large" as the second.

Example 2.2. For example, in the data presented in the Introduction to this chapter, if we computed the ratio of the number of homicides in New York to the number of homicides in Arkansas (both in 2016), we would obtain

$$\frac{335}{181} \doteq 1.85.$$

The 1.85 is a pure number which tells us that the raw number of homicides in New York was 1.85 times as large, or almost twice as large, as the number of homicides in Arkansas. We could also say that the number of homicides in New York City in 2016 was

$$\frac{335}{181} \times 100\% \doteq 185\%$$

of the number of homicides in Arkansas that year. △

Percentages also feature in the computation of *percentage changes* (sometimes phrased as *percentage differences*). These ideas are notoriously "slippery"[7] for many people, so please make sure that you understand them in detail, as we will be using

[7]That is, easy to get wrong and/or misunderstand.

them repeatedly throughout the course. The situation is always something like this: we measure some quantity twice, yielding an "old value" and an updated, "new value," or we consider an operation on the old value that changes it into the new value. The percentage change is a way to represent the difference between the two values. It is *the change as a percentage of the "old value."* To compute the percentage change, we always use

$$\boxed{\text{percentage change} = \frac{\text{new value} - \text{old value}}{\text{old value}} \times 100\%} \tag{2.3}$$

If the new value is larger than the old value, the quotient will be positive and the percent change will represent a percent increase; if the new value is smaller, we will have a percent decrease. I believe that the "slipperiness" mentioned before comes mainly from the fact that we need to identify which quantity represents the old value and which represents the new value when we apply the formula. This requires some "common sense" in many cases – you can't just fly blind; you need to think through which is which.

For instance, in (2.2), note that we treated the New York City per capita rate as the "old value" because we wanted to think of the Arkansas per capita rate as the "new value" and compare that to the New York City rate. It would be possible to make the comparison the other way too, but then the description would change and the formula would be used with the values reversed. The percentage difference

$$\frac{.000039 - .000063}{.000063} \times 100\% \doteq -38.1\%$$

would say that the New York City per capita homicide rate is about 38.1% smaller than the Arkansas rate. In the context of *percentage differences*, it may in fact help to think of the "old value" as a *reference value* and the "new value" as a *comparison value* that is being compared to the reference value.

We start with some everyday examples.

Example 2.3. If a man's suit that regularly sells for \$300 (the old value) is offered on sale for \$200 (the new value), this is a

$$\frac{200 - 300}{300} \times 100\% \doteq -33.3\%$$

percentage decrease or discount. On the other hand, if we knew the suit was being sold at a 25% discount, this should mean that the reduced price (the new value) is some number \$$x$ where

$$\frac{x - 300}{300} \times 100\% = -25\%.$$

If you do the algebra to solve for x (or even better, think: what is 25% of \$300? then subtract that from 300), you will see that the reduced price must be $x = \$225$. This is different from taking x plus 25% of x. We aren't taking 25% of the *old value (price)* that way!

Similarly, suppose the wholesale price of a refrigerator is \$500, but the appliance store sells it with a 20% markup. Now the old value is the wholesale price, and the retail price, which equals the wholesale price plus the markup, is the new value. The retail price must then be \$600, since

$$\frac{600 - 500}{500} \times 100\% = 20\%$$

gives the 20% increase. △

We saw another similar example in the Introduction too.

Example 2.4. Look back at (2.1) in the Introduction to this chapter. There the new value was the 2016 homicide total of 335 and the old value was the 2015 homicide total of 352 and the decrease of 17 homicides represented a percentage decrease of approximately 4.8%. △

Here is an example related to the Chapter Project from Chapter 1.

Example 2.5. The maximum depth of the Greenland ice sheet shown in Figure 1.3 from Chapter 1 is about 3205 meters. This comes from depth estimates made about 2013. Earlier estimates reported by Bamber, Layberry, and Gogineni[8] gave the maximum depth as 3367 meters. This estimate was based on observations from the 1970's through the 1990's, so it will serve as our "old value." We will take the "new value" to be the 3205 meters. This gives a percent change of

$$\frac{3205 - 3367}{3367} \times 100\% \doteq -4.8\%,$$

(coincidentally) another percentage decrease of about 4.8%.

Some care should be exercised in interpreting this result. This might represent an actual decrease in the thickness of the Greenland ice sheet. But it might also conceivably simply reflect more accurate measurements yielding smaller depth estimates. We have not presented enough information to make a firm determination of which of these alternatives is more probably correct. However, on the face of it, it seems likely that the thickness has actually decreased, given the fact that large melting events have been observed from the early 2000's to the present. △

A closely related idea is the *percentage relative error* in a measurement when an exact theoretical value is available:

$$\boxed{\text{percentage relative error} = \frac{\text{measured value} - \text{exact value}}{\text{exact value}} \times 100\%.} \tag{2.4}$$

This is arranged to give positive results when the measured value is larger than the exact value and negative results when it is less.

Example 2.6. Suppose an object's mass is exactly 34.7 kg, but we measure its mass to be 35 kg. The percentage relative error in this measurement is

$$\frac{35 - 34.7}{34.7} \times 100\% \doteq .86\%.$$

The absolute error (the actual amount by which the measurement differs from the exact value) in this case is .3 kg which is not an especially small mass. Nevertheless, we would say this was a rather *accurate* measurement in relative terms because the error is a very small fraction (less than 1%) of the exact value. △

So far we have been discussing ratios where the numerator and denominator have the same units. But ratios where the numerator and denominator have different associated units are also very common and useful. In this case the quotient value also comes with associated units that are something new and different from either the units of the numerator or the denominator. If the units of the numerator

[8] Journal of Geophysical Research, vol. 106, D24, December 27, 2001, pp. 33,773-33,780.

are N and the units of the denominator are D, then we call the units of the quotient "N per D," and write this as N/D. For instance, if the numerator was the total cost of a hotel stay in units of \$ and the denominator was a length of time in days, then the ratio would have units of \$/day, or dollars per day.

Example 2.7. As we said before, the population of New York City in 2016 was about 8.54×10^6 people. The city occupies a land area of about 302.6 square miles. The ratio

$$\frac{\text{population}}{\text{land area}} \doteq \frac{8.54 \times 10^6}{302.6} \doteq 28,222 \text{ people/square mile}$$

is called the (average) *population density* of the city. The meaning is this: If the total population was spread evenly over the whole area[9] then there would be over 28,000 people per square mile. This is an extraordinarily large population density, of course, showing the way multi-story apartment buildings and all the associated infrastructure of a large modern city can allow large numbers of people to live within a small geographic area. △

If you have studied velocities in a physics or calculus class, the following example will be familiar.

Example 2.8. If a moving object travels along a straight line path for a distance of $\Delta x = 34$ m in a time $\Delta t = 5$ sec, then we say it is moving at the (average) *velocity*

$$v_{ave} = \frac{\Delta x}{\Delta t} = \frac{34}{5} = 6.8 \text{ m/sec.}$$

The ratio $\frac{\Delta x}{\Delta t}$ computes *the distance traveled per second*, if the object moves at the same speed for the whole time (or the average distance traveled per second if not). Note that the units here are m/sec ("meters per second"). △

One purpose of ratios like the ones giving the average population density of cities is that they make it possible to compare different examples of cities on an "apples to apples" basis, independent of the raw total population figures. The population density can have a greater influence on the quality of life and "feel" of a city than the raw total population figure. For example, you will see in Exercise 6 that even though Chicago has a much larger total population than Boston, that population is not as densely packed because the total area of Chicago is also much, much larger.

2.3. Concentrations

Environmental scientists often use ratios together with a variation on the notion of percents to describe the *concentrations* of pollutants or contaminants in air, water, foods, etc. The idea is described in the following example.

Example 2.9. Burning of coal and some other industrial processes release small amounts of mercury and other heavy metals into the atmosphere. That mercury is then incorporated into the bodies of fish and land animals as they eat plants, smaller fish, and smaller animals that have been contaminated with it. As

[9]This is not actually the case. Manhattan is more densely populated than the four other New York City boroughs. See Exercise 7 below.

a result, animals at the tops of food chains tend to concentrate mercury in their tissues. Mercury levels in fish are of concern for human health because fish is an increasingly important part of many human diets and ingesting even very small amounts of the metal can damage the central and peripheral nervous systems.

According to the U.S. Food and Drug Administration[10], the flesh of the following popular shellfish and fish caught for human consumption as food (data collected through 2012) contained mercury at the levels shown in Table 1. The units used

Table 1. Mercury Concentrations in Fish (in parts per million, ppm)

Species	Concentration in ppm
Scallop	.003
Shrimp	.009
Salmon (canned)	.014
Salmon (fresh, frozen)	.022
Catfish	.024
Lobster	.107
Tuna (canned, light)	.126
Tuna (canned, white)	.350
Grouper	.448
Swordfish	.995

here (parts per million by mass) represent the contaminant (the mercury) as a fraction of the total mass of the fish. This gives a common way to represent very small amounts without using extremely small mass or volume values. Those very small values might cause confusion and misunderstanding.

What this means, for instance, is that flesh from a type of fish would contain 1 part per million mercury by mass if a $1,000,000$ g sample of the fish contained 1 g of mercury:

$$\frac{1 \text{ gram mercury}}{1,000,000 \text{ grams fish}} = 1 \text{ ppm}.$$

The swordfish mercury level in the table above is nearly that high; the grouper mercury level as about one half that high; and the catfish mercury level is about 1/40 times that high.

The values in Table 1 might seem extremely small in everyday terms and we might be tempted to think that there is nothing to worry about. The bad news is that it takes very little mercury to produce the nervous system damage we mentioned before. In fact, on the basis of this data, the Environmental Protection Agency recommends that children (who are more susceptible to mercury poisoning

[10] `https://www.fda.gov/food/foodborneillnesscontaminants/metals/ucm115644.htm`, consulted June 23, 2017.

because of their smaller body weights and developing nervous tissue) eat white (albacore) tuna no more than *once a month* based on this level of mercury! Adults are recommended to limit their intake of white (albacore) tuna to about *three times a month.* △

Larger concentrations could also be reported using units of parts per thousand; smaller amounts might be reported using units of parts per billion. The idea is the same.

2.4. Probabilities

One of the commonly-used conceptual frameworks for thinking about probabilities, called the *frequentist approach*, uses ratios to define and compute probabilities.[11]

For the purposes of this introductory discussion, we will stick to situations where the total number of possible outcomes of a measurement or an experiment is a *finite set* called the *sample space*, S. It is often the case that each single element of the sample space is considered to be equally likely as an outcome. If so, then we assign a *probability* of

$$\frac{1}{\text{number of elements of } S}$$

to each element of the sample space. Note that each of these values is ≥ 0 and adding up all of them (i.e for each element of S) gives a total of 1. If the elements of the sample space are not all equally likely as outcomes, then each would be assigned some probability ≥ 0, and these would again sum to a total value of 1. Either way, we have what is called a *probability space*, or *probability model* consisting of the sample space and the assignment of probabilities.

Example 2.10. Suppose that our experiment consists of selecting one randomly chosen fish out of a pool containing 13 carp, 10 sunfish and 3 northern pike. If each fish is equally likely as the one selected, then the probability of picking any one individual fish would be $\dfrac{1}{13+10+3} = \dfrac{1}{26}$. △

An *event* is a subset E of the sample space. Using the assignment of probabilities to the elements of S, we define the probability of E, written $P(E)$, as *the sum*[12] *of the probabilities of the elements of* E. In the equal-probability case, this gives

$$P(E) = \frac{\text{number of elements of } E}{\text{number of elements of } S} = \frac{\text{outcomes in } E}{\text{outcomes in } S}. \tag{2.5}$$

This should become more understandable if we pause for some examples to illustrate these ideas.

[11] There is also an alternative school of thought in probability and statistics known as the *Bayesian* approach, where probabilities are treated as estimates and updated on the basis of the information supplied by experiments. This is increasingly important in scientific applications and we may be able to discuss some of the basics of Bayesian probability in the spring semester.

[12] Technical note: Recall that we are restricting to the case where S is a finite set. If S is an interval on the real line or a region in the plane or in a higher-dimensional space, then probabilities are often described by a corresponding *probability density function* and techniques from integral calculus can be used to compute them. Many Calculus 2 courses study those applications of integration to computing probabilities.

EXAMPLE 2.11. Continuing from Example 2.10, suppose that E represents the outcome that we select a sunfish. Since each of the fish is equally likely as the one selected and there are 10 of the sunfish, the probability of E would be obtained by adding together 10 terms equal to $\frac{1}{26}$:

$$P(E) = \frac{1}{26} + \cdots + \frac{1}{26} = \frac{10}{26} = \frac{5}{13} \doteq .385. \quad \triangle$$

Here is another, somewhat more involved, example.

EXAMPLE 2.12. Suppose you roll two cubical (six-sided) dice, one with black dots on a white background, the other with white dots on a black background. Each of the dice will come up with one face facing upward, showing a number of dots between 1 and 6. Because the two dice are visually different, note that we can choose to record the results by showing the number from the die with black dots first, then the number from the die with white dots second. This means that in mathematical terms, the possible outcomes of one roll are in one-to-one correspondence with the *ordered pairs* (m, n) where m is $1, 2, 3, 4, 5$, or 6 and similarly for n. This is the sample space S for the experiment of rolling the two dice once, and we could list all of its $6 \times 6 = 36$ elements like this:

$$S = \{(1,1),(1,2),\ldots,(1,6),(2,1),(2,2),\ldots,(2,6),\ldots,(6,1),(6,2),\ldots,(6,6)\}.$$

In this simple example, if the two dice are both "fair," that is, if they are perfectly balanced and not "loaded" so that some numbers come up more frequently than others, then it would make sense to assign the same probability $\frac{1}{36}$ to each of the elements of S (the possible outcomes of the roll).

On the other hand, suppose the two dice were both loaded in the same way so that 3's came up $\frac{1}{5}$ of the time but each of the other numbers came up $\frac{4}{25}$ of the time. Note that

$$\frac{1}{5} + 5 \times \frac{4}{25} = \frac{1}{5} + \frac{4}{5} = 1,$$

as we required above. Then the probability of rolling a 3 on both dice is

$$P((3,3)) = \frac{1}{5} \times \frac{1}{5} = \frac{1}{25} = 0.04. \tag{2.6}$$

The probability of rolling one 3 and something else (that is, $(3, n)$ with $n \neq 3$ or $(m, 3)$ with $m \neq 3$) is:

$$P((3,n)) = P((m,3)) = \frac{1}{5} \times \frac{4}{25} = \frac{4}{125} = 0.032. \tag{2.7}$$

Finally, if m, n are both different from 3,

$$P((m,n)) = \frac{4}{25} \times \frac{4}{25} = \frac{16}{625} = .0256. \tag{2.8}$$

We see that rolling two threes is more likely than either of the other two possibilities.

Now, let us consider the event E described in words by saying that the die with black dots shows an *even number*, while the die with white dots shows an *odd number*. As a subset of S,

$$E = \{(2,1),(2,3),(2,5),(4,1),(4,3),(4,5),(6,1),(6,3),(6,5)\}.$$

Hence if the dice are not loaded, we will have

$$P(E) = \frac{\text{number of elements of } E}{\text{number of elements of } S} = \frac{9}{36} = \frac{1}{4}.$$

In the exercises, you will compute $P(E)$ if the dice are "loaded" as above. △

The frequentist interpretation of the *meaning of these probabilities* is that if we repeated the experiment for a large number of trials and kept track of the ratio

$$\frac{\text{number of outcomes in } E}{\text{number of trials}}, \tag{2.9}$$

then this ratio would tend to approach $P(E)$ as the number of trials increased without bound. This statement is often called the *Law of Large Numbers*. For example, with the two fair dice as described in Example 2.12, if we roll the dice a large number of times, then roughly $\frac{1}{4} \times 100\% = 25\%$ of the rolls should show an even number of black dots and an odd number of white dots. If the rolls are truly random, there could be more or less than 25% of them meeting this condition. However, the deviation (measured as a percent difference) from value of 25% should tend to decrease the more times we roll the dice. Moreover, in the absence of a theoretical probability model as we described above, a frequentist might take a value computed as in (2.9) as an "experimental probability" or estimate of the unknown theoretical probability.

2.5. Proportions

Mathematically, we say a *proportion* is the statement that two ratios are the same:

$$\frac{A}{B} = \frac{C}{D}.$$

Alternatively, we might say A has the same proportion to B as C has to D if this equation is true. Note that if we know *three* of the terms in a proportion, then it is always possible to solve for the remaining term by simple algebraic manipulations.

Example 2.13. Any time we estimate a distance from a map and the associated distance scale we are doing something similar. If the map says 1.5 inch = 10 miles, and we measure a distance of 2.5 inches on the map, then the real world distance x between those two points will satisfy the proportion

$$\frac{x}{2.5} = \frac{10}{1.5},$$

so $x = 2.5 \times \frac{10}{1.5} = 37.5$ miles. Similarly, if we were constructing a map with the scale 1.5 inch = 10 miles and we knew the actual distance between two locations to be 45.3 miles, then the separation between the points on the map representing those locations would be x satisfying the proportion

$$\frac{45.3}{x} = \frac{10}{1.5},$$

so

$$x = \frac{45.3 \times 1.5}{10} \doteq 6.8 \text{ inches}.$$

The conversions from one system of units to another from Chapter 1 and the calculations of distances and areas we were doing in the Chapter Project were based on proportions. △

The idea of proportionality is also used in the following way. We say two quantities A and B *are proportional*, or equivalently A *is proportional to* B, if there is some constant c such that $A = cB$. The constant c is called the *proportionality factor*. This means that if we know the value of B, then the corresponding value of A is obtained by simply multiplying by the constant c.

Example 2.14. The area of a circle is proportional to the square of its radius and also to the square of its diameter.[13] This true since if r is the radius, we have the formula $A = \pi r^2$, and $\pi \doteq 3.14159$ is constant. Moreover, if d is the diameter then $r = d/2$, so $A = \frac{\pi}{4} d^2$, with proportionality factor $\pi/4$. $\triangle$

Capture-Recapture. Proportions are used in a clever way by biologists and environmental scientists when they want to estimate the population of some species of animals in a habitat.[14] The motivation for this method is that it is usually not practical to do an exhaustive census of the whole population. It may be too difficult or time-consuming to capture and count all of the animals present. However, it *is often possible* to capture and tag a random sample of the animals in the population, then release them back into the population and let them "mix in" with the rest. After enough time is allowed for mixing (but not too much time, so that not many births and deaths take place), a new sample is taken and the proportion of tagged animals in the new sample is computed. Assuming that the animals that were tagged from the first sample were mixed thoroughly into the population before the second sample was taken we would expect the following proportion to hold (at least approximately):

$$\boxed{\frac{\text{size of first sample}}{\text{total population}} = \frac{\text{tagged animals in second sample}}{\text{size of second sample}}.} \tag{2.10}$$

Three of the terms here, namely all except for the total population, are known. This means we can use the proportion to estimate the total population. This process is called the *capture-recapture, or catch-and-release* method.

Example 2.15. Suppose a first sample of 25 muskrats are captured, tagged and released in a forest. A week later, a new sample of 40 muskrats is taken, and 6 of them are found to be animals tagged in the first sample. We want to use this information to estimate the population of muskrats in the forest. Using (2.10), and writing x for the total population, which is the unknown quantity, we see

$$\frac{25}{x} = \frac{6}{40} \Rightarrow x = \frac{25 \cdot 40}{6} \doteq 167$$

muskrats. $\triangle$

2.6. Chapter Project

The project for this chapter has two goals:

---

[13]For those with an interest in the history of mathematics, this is the way areas of circles were treated in many ancient Greek texts, for example in Proposition 2 of Book XII of Euclid's *Elements*. Euclid does not mention the proportionality factor π because he did not use numbers as measures of areas. The value of π was later approximated rather closely by Archimedes.

[14]The method works best when the habitat is *closed* so there is no in- or out-migration of the animals.

(1) To continue our work with ratios, percentages, and proportions on some world population and GDP data from the U.N., and
(2) To introduce the Excel spreadsheet program[15] that we will be using extensively for the rest of the year.

The following instructions will lead you through a first practice Excel session working with a spreadsheet file that you will download from the course homepage.

Getting Started with Excel. Launch Excel and take a look at the overall layout of the of the window. There are tabs, menus, etc. similar to many standard programs, but there are some differences too. In particular if you are using the Windows Excel 2007 or later version, note the large "Office Button" at the upper left. This is where all of the usual File options are located (i.e. the controls for reading in or saving files, printing, etc.)

Like all spreadsheet programs, Excel gives you a workspace that is composed of a 2D grid of "cells" identified by location – by an *address*. The columns are labeled by capital letters, and the rows are labeled by numbers.

- A single cell is referenced by the column, followed by the row, for instance `B23` is the cell in column B and row 23.
- A range of cells is referenced by giving the "starting cell," a colon, and the "final cell" in the range. For instance `B2:B45` indicates the cells in column B and rows 2 through 45. `B2:F2` indicates the cells in row 2 and columns B through F. Similarly, `B2:D10` indicates all the cells in a *rectangular block* with upper left corner at cell `B2` and lower right corner at cell `D10`.
- The addresses seen so far are all *relative addresses.* In other words, they are set up so that if we perform an operation in one cell that depends on the entries to the left in its row, then it is possible to copy and paste that operation to other rows and the entries in the new row will be used. If you want to specify a *fixed* address then put in $ characters: `$C$5` means the cell with fixed address in column C and row 5. (We will see several examples of this in a while; if it is not clear why we need this distinction, wait until you see the examples!)

The contents of a cell can be a text label identifying what the data in a row or column represents, a number, or a formula indicating how to perform a desired calculation using other information in different cells within the spreadsheet. When you finish entering a formula this way and press the Enter key, the indicated computation is performed and the result is displayed in that cell. One *very nice* feature of spreadsheets is that if you change the contents of a cell that is used to compute a value this way, then the calculation is automatically performed again to update the value displayed. We will also see this in a moment.

A First Worked Example. Begin by reading in the spreadsheet `First.xls` from the course homepage. (If you have not done so already, download and save that in a location where you can have Excel read it in.)

- Press the "Office Button" at the upper left of the Excel window (or use the File pull-down menu in Mac versions),
- then Open,

[15]Google spreadsheets and freely available software such as the LibreOffice suite can also be used here.

- Find the appropriate location where you saved the `First.xls` file in the folder box at the top of the Open window, highlight the file `First.xls`,
- and press Open at the bottom.

You should now see a rectangular block of cells filled with names, text, and numbers at the upper left of the spreadsheet in rows 1 through 10 and columns A through E. Think of this as the grade book for a small class with 8 students (the rows are labeled with their names) who have had four assignments as in the labels for columns B through E. Note that `A12` has the text "Average" but there are no numbers on that row (yet!). We are going to use Excel to compute the averages on each assignment.

- In cell `B13`, enter the formula =`AVERAGE(B2:B9)`. As you type, you will see this showing up in the cell and in the input box above the grid. When you are done press Enter, and the average will be computed and displayed.
- Now we will use the same method to compute the average on each of the other assignments: Highlight cell `B13` by clicking the left mouse button over that cell. Make sure the Home tab at the top of the Excel window is active, press Copy (next to the "Office Button"), drag the highlighting box so that all the cells in row 13, columns B to E are highlighted, and press Paste (next to Copy). You should now see the averages for each column.
- In doing the averages we were making use of the *relative addressing* mentioned above. Copying the formula in one cell and pasting it into another also changed the addresses of the cells that the formula was applied to. Now, we are going to perform an operation where we want to use contents of a fixed cell on multiple rows. Start by filling in new information in row 14: Put a text label "Weights" in `A14` and the constants .3 in `B14`, .25 in `C14`, .4 in `D14`, and .05 in `E14`.
- In cell `F1` add the text label "Course Average." In `F2` enter the formula
  ```
  =$B$14*B2 + $C$14*C2 + $D$14*D2 + $E$14*E2
  ```
  You should see the weighted average displayed.
- You can now copy and paste that formula to the other cells in column F and rows 3 through 9 to do the same computation for the other students in the class. (Note that the weights always come from the same row, hence the fixed addresses. Can you see what would happen if we did not do it that way?)

Here is some other useful information:

- There are a number of standard mathematical functions that can be applied to numerical contents of spreadsheet grids. In an Excel formula you can square the contents of a cell by saying, for instance `B13^2`.
- If you want to take the square root of something computed from information other cells, you use `SQRT()`. For instance, to compute the square root of the sum of `B13^2` and `C13^2`, you could enter
  ```
  = SQRT(B13^2 + C13^2)
  ```
  in another cell of the spreadsheet. When you press Enter, the value will be computed and displayed in that cell.

The Data. The country-by-country population information in Table 2 (on the next page) comes from Table 3 of the 2015 United Nations Demographic Year-book[16]. Notes:

(1) Only a selection of the roughly 200 nations are listed(!) The first seven nations are located in Africa, the next seven are in North and South America (including the Caribbean), the seven following those are in Europe, and the last seven are in Oceania and Asia.
(2) The male and female populations are according to the most recent available national census (which varies by year according to the nation), and are in units of 1000s.
(3) The 2010 and 2015 midyear populations are estimates given in units of 1000s
(4) The surface area is in square km.
(5) Unavailable data is marked by a $*$
(6) GDP is the market value of all final goods and services produced in a year. The given figures are estimates for 2015, in units of trillions of U.S. dollars.

Questions. Create a spreadsheet, enter this data, then make the following computations. Answer the questions below in a separate text file or document.

(A) For each of these 28 nations, compute the male population as a percentage of the total population *according to the most recent national census*. In which nations is the male percentage greatest? In which is it smallest? Why does it work out this way?
(B) For each of the 28 nations, compute the percentage difference between the male and female populations.
(C) Compute the *percentage change* in the populations of each of these nations between 2010 and 2015. (You won't be able to do this in two cases because of missing data; just ignore those nations for this question.) How many of these nations are losing population, and which has the largest percentage decrease? Which has the largest percentage increase?
(D) Compute the population densities in people per square kilometer for each of these nations. Use the 2010 mid-year estimates, since some data is unavailable for 2015. Which nations have the largest population densities among the 28 listed here? (Note: the highest population densities actually tend to occur for smaller, especially island, nations).
(E) What are the 5 richest and 5 poorest of these countries, and how should you make a fair comparison?
(F) Recall that the GDP figures are estimates for 2015. Assuming that the per capita GDP was unchanged from 2010 to 2015 (that is that the proportion of GDP to population was the same both years), estimate the GDP figures for each of the countries in 2010.
(G) What proportion of the total estimated 2010 population of the world, 6.9 billion, is accounted for in the 28 countries listed in Table 2?

[16] `https://unstats.un.org/unsd/demographic/products/dyb/dybsets/2015.pdf`, accessed on June 26, 2017.

Table 2. Extracts from 2015 U.N. Demographic Yearbook and 2015 GDP Figures

Nation	Male	Female	2010 Mid.	2015 Mid.	Area	GDP
Egypt	37,219	35,579	78,685	88,958	1,002,000	.282
Kenya	19,192	19,418	40,406	45,509	591,958	.061
Mozambique	9,747	10,506	22,417	25,728	799,380	.017
Namibia	1,022	1,091	2,143	2,281	824,116	.013
Niger	8,519	8,620	15,204	19,125	1,267,000	.008
Nigeria	71,345	69,086	159,619	*	923,768	.568
Senegal	6,428	6,445	12,509	14,357	196,712	.016
Argentina	19,524	20,593	40,788	43,137	2,780,400	.543
Brazil	93,407	97,349	195,498	204,451	8,515,767	1.77
Canada	16,414	17,062	34,005	35,849	9,984,670	1.55
Chile	7,448	7,669	17,094	18,006	756,102	.258
Dominican Rep.	4,739	4,706	9,479	9,980	48,671	.064
Honduras	4,052	4,251	8,046	8,577	112,492	.019
Mexico	54,855	57,481	114,256	121,006	1,964,375	1.14
U.S.	151,781	156,964	309,347	321,419	9,833,517	18.03
Austria	4,094	4,308	8,361	8,576	83,871	.437
Croatia	2,066	2,219	4,295	4,225	56,594	.057
France	29,715	31,685	62,918	64,395	551,500	2.42
Germany	39,146	41,074	81,757	81,198	357,376	3.36
Greece	5,303	5,513	11,121	10,858	131,957	.238
Poland	18,420	19,624	38,517	38,006	312,679	.545
U.K.	31,126	32,254	62,759	64,875	242,495	2.86
Armenia	1,347	1,525	3,256	3,011	29,743	.011
Australia	10,737	10,990	22,032	23,778	7,692,024	1.23
China	686,853	652,872	1,337,700	1,371,220	9,600,000	11.16
India	623,270	587,585	1,182,105	*	3,287,263	2.12
Iran	37,906	37,244	74,340	78,773	1,628,750	.425
Japan	61,829	65,281	128,070	126,958	377,930	4.38
Pakistan	67,840	62,739	173,510	191,710	796,095	.251
Saudi Arabia	15,307	11,830	27,563	31,016	2,206,714	.653

Chapter Exercises

(1) In the Chapter Project for Chapter 1, we said that the freshwater ice in the Greenland ice sheet constituted about $\frac{1}{9}$ of the world's freshwater ice. What percentage of the world's freshwater ice is this?

(2) The total population of the U.S. in 2010 was $308,745,538$ according to the Federal Census. The 2017 population is estimated to be $324,700,000$.

  (a) What was the percentage change in the population from 2010 to 2017?

  (b) What was the percentage change per year?

(3) This exercise investigates the concept of the percentage relative error from (2.4).

  (a) Using a meter stick, a student measures the distance from the floor to the ceiling in a room and finds a distance of 3.44 meters. The exact floor to ceiling distance is known to be 3.40 meters. A second student measures the height of a large book using a ruler and finds a length of .44 meters. The actual height is .40 meters. What are the *absolute* and *percentage relative errors* of these measurements?

  (b) Which student was more *accurate* in the measurement?

  (c) State in general how you might use the absolute and/or percentage relative errors of measurements to assess their accuracy.

  (d) Is *precision* of a set of measurements the same as *accuracy*? Explain with an example. (Note: You may need to look this up. If you do, please cite your source!)

(4) From the web site `www.campusexplorer.com`, accessed June 22, 2017:

> "The ratio of men to women in college has been slowly and steadily shifting for decades. More women have enrolled in college since 1979 and the ratio seems to have only recently stabilized at around 57%. That means that men make up only 43% of the national student population."

  (a) Does this statement make sense mathematically? Criticize from the point of view of the concepts introduced in this chapter.

  (b) What should the correct value of the ratio of men to women be if, as the authors apparently intended to say, women make up 57% of the student body and men make up 43% of the student body?

(5) According to data from the Solar Energy Industries Association[17], the "top 10 solar states" in terms of (cumulative) total solar electric generating capacity (measured in megawatts, Mw) installed in solar panel farms, homes, etc. as of 2016 are given in Table 3.

  (a) Look up estimates for the populations of these states in 2016, fill in that column of the table, then use that data to compute the *per capita solar electric capacity* for each state and fill in the last column.

  (b) Discuss any interesting patterns you see here. In particular, if you make the comparison on the per capita basis, are the states ordered the same way or do some states move higher or lower in the rankings?

(6) In Table 4, we give the populations and geographic areas of some of the larger U.S. cities according to the 2010 Federal Census. Notes: The populations and

[17] `http://www.seia.org/research-resources/top-10-solar-states`, accessed on June 22, 2017.

Table 3. Table of Solar Electric Generating Capacity by State

State	Capacity in Mw	Population	Capacity Per Capita
California	18,296		
North Carolina	3,016		
Arizona	2,982		
Nevada	2,191		
New Jersey	1,991		
Utah	1,489		
Massachusetts	1,487		
Georgia	1,432		
Texas	1,215		
New York	927		

areas are given only for the incorporated cities, not the surrounding metropolitan areas. The figures for New York City are different from those given in the text because those were estimates for the year 2016.

Table 4. Table of City Populations and Areas in 2010

City	Population	Land Area (in square miles)
Boston	645,149	48.43
Chicago	2,695,598	227.13
Miami	399,457	35.67
New York City	8,175,133	302.6
Philadelphia	1,517,550	135.09
San Francisco	805,235	46.69

(a) Compute the average population densities for all of these cities in units of people/square mile, then people/square kilometer.

(b) What was the *percentage change* in the average population density of New York City between 2010 and 2016? Does it make any difference which set of population density figures (i.e. the two in units of people/square mile or the two in units of people/square kilometer) you use here? Explain.

(7) In Table 5, we give the populations and land areas of the five boroughs of New York City, again according to the 2010 Federal Census.[18]

[18] You should check that they add up to the total given for the whole city in Exercise 6. The total reported for Manhattan on the Wikipedia page

TABLE 5. Table of New York City Borough Populations and Areas in 2010

Borough	Population	Land Area (in square miles)
Bronx	1,385,108	42.10
Brooklyn	2,504,700	70.82
Manhattan	1,585,873	22.83
Queens	2,230,722	108.53
Staten Island	468,730	58.37

(a) Compute the average population densities for all of these in units of people per square mile, then people per square kilometer.

(b) Is the numerical average of the 5 borough population densities the same as the overall New York City population density from Exercise 6? Why or why not? If you knew the five borough population densities from this problem, what could you do to recover the city population density?

(8) Refer to Table 1 in Example 2.9.

(a) How much mercury (in grams) would we expect to find in a 25 kg sample of grouper flesh?

(b) What is the percentage difference between the mercury level reported for white (albacore) tuna and light (skipjack) tuna? Based on that proportion, how often do you expect the EPA recommends that children and adults can eat light tuna? Check this online.

(c) What is the percentage difference between the mercury level reported for swordfish and shrimp? Why should it be true that mercury levels are so much higher in swordfish than in shrimp? If you need to, look up the feeding habits of these animals.

(d) What is the percentage difference between the mercury level reported for lobster and scallops? Why should it be true that mercury levels are so much higher in lobster than in scallops? If you need to, look up the feeding habits of these animals.

(9) (Hypothetical; this is not real data.) A scientist sends some samples of water contaminated with lead to a commercial lab for analysis. To check the accuracy of its measurements, the lab first does a calibration run on a sample known to have a concentration of 150 ppb (parts per billion). The lab's machine returns a measured concentation of 159 ppb.

(a) What is the percentage relative error in this measured concentration?

`https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population_density` is different for some reason that I do not understand. The figure there is either a simple transcription error, or it may also not include populations like that of Roosevelt Island, which is administratively part of the borough of Manhattan but not contiguous with Manhattan Island.

(b) One of the scientist's samples yields a value of 430 ppb when the lab analyzes it. Assuming that measurement has the same percentage relative error as the calibration run, estimate the actual lead concentration in ppb in that sample.

(c) Discuss whether the procedure described in part (b) is necessarily justified. (For instance, what if the machine was just mis-calibrated and all the values it returned were biased by the same absolute amount?)

(10) Refer to Example 2.12.

(a) Show that the probabilities for the "loaded" dice given in (2.6), (2.7), and (2.8) add up to 1 if we sum over all the elements of S. (Hint: You will need to determine how many outcomes there are in each case.)

(b) If all pairs (m, n) are equally likely, what is the probability of the event $E =$ the numbers of dots sum to 7?

(c) What is the probability of the event $E =$ number of black dots even and number of white dots odd in the loaded case from part (a)?

(11) The severity of hurricanes, floods and other disasters are often described in terms related to probabilities. For instance a "100-year hurricane" is a hurricane that on average would only be expected to occur (only) once in a 100-year period, while the other hurricanes in that period are less severe.

(a) How many 100-year hurricanes would be expected to occur in a 500-year period? In a 50-year period? In a 1-year period? Explain. (Hint: numbers between 0 and 1 interpreted as probabilities are OK.)

(b) Which would be expected to be more destructive – a 100-year hurricane or a 500-year hurricane? Explain.

(12) Express each of the following statements as an equation by using proportionality. (See Example 2.14.)

(a) The volume of a sphere is proportional to the cube of its radius.

(b) The volume of a right circular cylinder is proportional to the product of the square of the radius and the height.

(c) The volume of a right circular cone is proportional to the product of the square of the radius and the height.

(d) "Newton's Gravitation Law:" The gravitational force exerted by one body on another is proportional to the reciprocal of the square of the distance between them. (The proportionality factor depends on the two masses.)

(13) Refer to (2.10).

(a) Give another explanation why this method works (approximately) by interpreting the ratios as probabilities.

(b) Which situation gives a *larger estimate for the total population*:

- a large proportion of the individuals in the second sample are found to be tagged, *or*

- a small proportion of the individuals in the second sample are found to be tagged?

Explain.

(c) A random sample of 500 mosquitoes in a neighborhood is captured, tagged, and released. A few hours later, a second sample of 400 is taken and the sample contains 10 of the tagged mosquitoes. What is the estimated population of mosquitoes?

CHAPTER 3

Part I Summary Project

To solidify your confidence in dealing with quantitative information, this chapter presents another project summarizing ideas developed in the previous two chapters, as well as a few data visualization topics that we have not mentioned before. The subject will be the whole U.S. energy economy in recent years.

3.1. Background

Figure 3.1 shows estimated energy production and end-use consumption data for all major sectors in the United States energy economy for the year 2010. Figure 3.2, given later, shows the analogous data for the year 2016. Similar diagrams have been produced each year by the Lawrence Livermore National Laboratory, using data provided by the Energy Information Administration, a division of the U.S. Department of Energy.

Reading the Charts. To read these diagrams, you will want to rotate the page or the screen by 90° clockwise, then follow the flow of energy by reading from left to right. On the far left are boxes representing the major energy sources labeled Solar, Nuclear, Hydro, Wind, etc. The 4 boxes towards the right (Residential, Commercial, Industrial, Transportation) are the four largest end-use sectors of energy consumption. Some of the basic source energy is first converted to electricity before it is transmitted to the four end-use sectors, as shown by the box labeled "Electricity Generation."[1] On the far right two boxes in the diagram indicate the total amount of energy that is lost or wasted ("Rejected Energy") and the amount that is used for its intended purpose ("Energy Services"). The "pipelines" joining the boxes represent how much energy originating in the box to the left was delivered to the box on the right. The numerical amounts are given as the numbers printed next to the pipelines in units to be discussed below.[2] The width of each energy flow (pipeline) is also in direct proportion to the amount of energy in that flow so you can judge the importance visually.

Units. The header on the top of the chart states that Net Primary Resource Consumption is approximately 98 Quads – this is energy-speak for 98 quadrillion BTUs. One quadrillion is 1×10^{15}. All energy measurements on the diagram have units of quadrillion British thermal units (BTU's), or "Quads" for short. One BTU is the quantity of heat needed to raise the temperature of 1 pound of

[1] At the top of the diagram, you can also see that there is a very small amount of electric energy that was imported (mostly from hydroelectric plants in Canada). This is so small that it is essentially negligible in the big picture.

[2] The numbers aren't always near the source, so you may need to hunt for them!

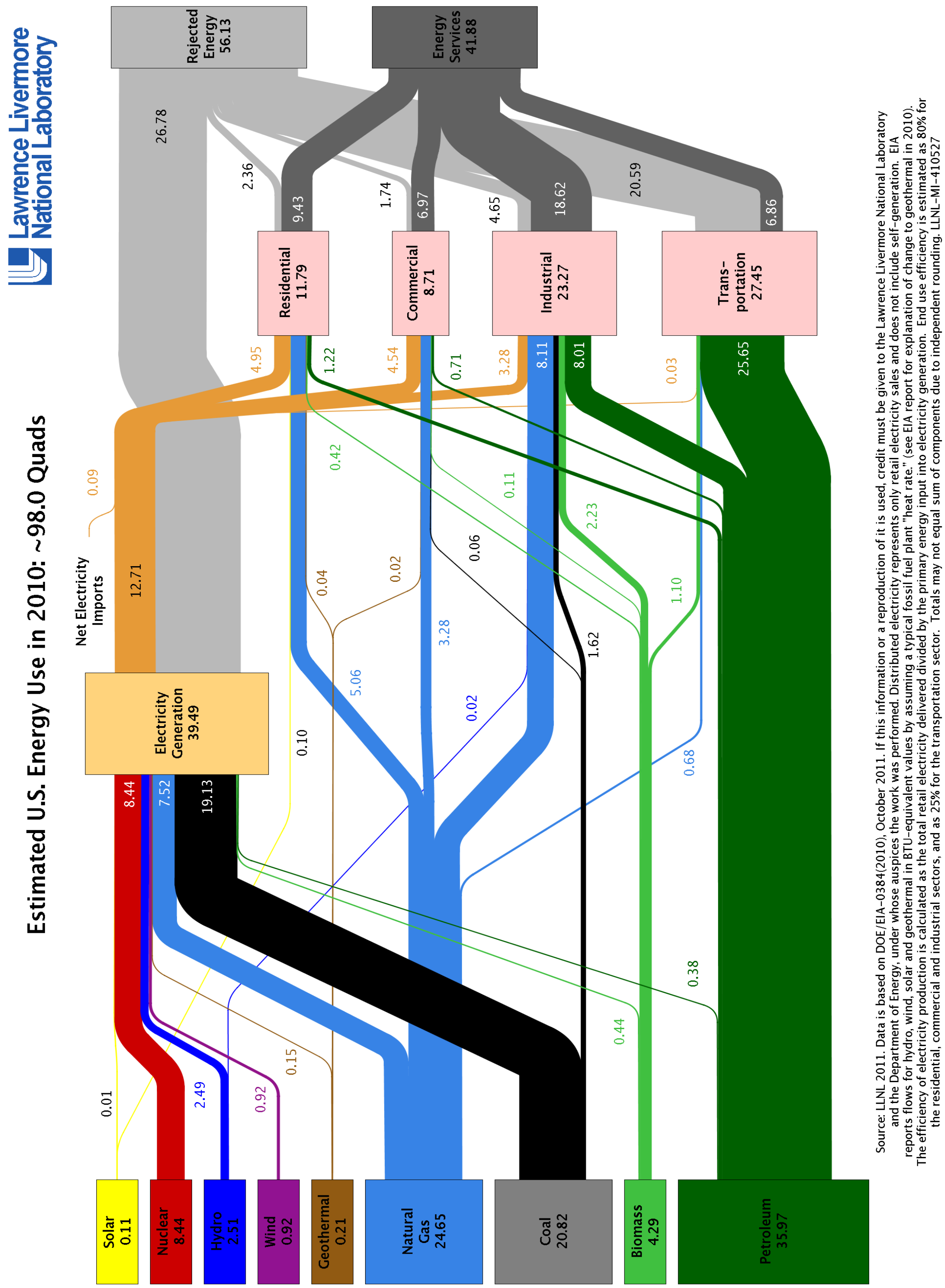


Figure 3.1. The U.S. energy sector in 2010.

water by 1° F at or near 39.2° F. Some approximate conversion factors are given below. In most cases, the numbers on the chart have been rounded to 1 decimal

TABLE 1. Energy content of various fuels.

Fuel	Equivalent (BTU)
1 ton coal	21,400,000
1 barrel (42 gal.) oil	5,800,000
1 cubic foot natural gas	1000
1 kilowatt − hour electricity	3,400

place (tenth of a Quad). Because of this rounding, the total Quads listed next to each production sector and end-use sector might not exactly equal the sum of the individual components.

3.2. Questions

(A) Rank the energy sources from highest to lowest in Quads and compute the percentage each accounts for in the total U.S. energy sector. Construct a *pie chart* representing this information. (You can do this in Excel, for instance, or construct the pie chart by hand.)

(B) Using Table 1, determine the answers to the following.

(1) What is the equivalent amount of petroleum used in 2010 in units of barrels?

(2) What is the equivalent amount of natural gas used in 2010 in cubic feet?

(C) Construct a table showing the total natural gas energy used by each of the four end-use sectors. Note that some of the natural gas is used for Electricity Generation, and that electricity is then used in the end-use sectors. In other words, do not miss the portion of the end-use sectors that use natural gas by way of electricity generated by burning gas. (How will you account for the fact that only a part of the Electricity Generation is done by burning natural gas?)

(D) Which of the energy sources are based on "fossil fuels?" Which of the energy sources are "renewable energy" sources? What percentage of the total energy produced is accounted for by renewables?

(E) **Analysis of the Electricity Generation sector.** Many sources of energy flow into the electric power sector, which then distributes electricity to the end-use sectors. Petroleum, coal, natural gas and biomass are burned in conventional power plants to produce heat to boil water. The steam from the boiling water spins turbines which then produce electricity. Nuclear fuels can be used to produce electricity in much the same way: nuclear reactions in power plants make the heat which boils the water which spins the turbines which produces the electricity. Other sorts of energy are also used to generate electricity. Both conventional-electric and nuclear-electric power plants have the property that a large amount of the fuel energy is lost in the process of making electricity and then more is lost during transmission along electrical lines.

(1) The 2010 diagram indicates that the electric power sector converted various energies to 39.49 Quads of electrical energy. What were the top two sources of energy for the electric power sector?

(2) How many Quads of electricity were successfully distributed from power plants, and how many Quads were lost at the power plants?

(3) The *efficiency* of an energy system is defined as the percentage of the total energy used for the intended purpose. Determine the efficiency of the U.S. electric power sector, using your previous answers and ignoring losses after distribution.

(4) Natural gas contributed 7.52 Quads of energy to the electric power sector. How many Quads of that contribution were immediately lost by the electric power sector? Explain any assumptions you are making.

(5) Give two practical reasons why so little electricity is distributed to the transportation sector.[3]

(F) Make a pie chart illustrating the *five* largest energy forms used in the Residential sector and the percentages they account for in the Residential sector.

(G) **The Industrial sector.** The Industrial sector includes manufacturing industries, mining, construction, agriculture, fisheries and forestry.

(1) The industrial sector consumed 23.27 Quads through *six* forms of energy.[4] Draw a bar chart to illustrate the number of Quads used for each of those six form of energy. Label each bar with its energy name and amount.

(2) Natural gas energy is directly consumed by the Industrial sector through the burning of natural gas. But the Industrial sector also consumes natural gas energy indirectly by using distributed electricity. Compute the total Quads of natural gas energy consumed by the industrial sector. Ignore energy losses.

(H) **Petroleum and Transportation.**

(1) The transportation sector is primarily fed by the energy derived from Petroleum, with small contributions from Biomass, Natural gas, and Electricity. What percent of the total Petroleum energy is consumed by the transportation sector? You can ignore the small amount of oil energy that is first converted to electricity.

---

[3] There is *some* electric power used for transportation, though. The Amtrak Northeast Corridor passenger rail lines between Boston and Washington, DC use electric power, for instance.

[4] You will need to look carefully to find them all, but they are there!

(2) What percent of the energy consumed by the transportation sector was wasted in 2010? Consider all forms of energy.

(3) Multiply the percentages found in the last two parts to find the percentage of all Petroleum energy that was wasted by the transportation sector. How many barrels of oil is that? How many gallons?

(I) **Heating homes.** Most homes and apartments today are heated with electricity or natural gas. (The exception to this general rule comes in New England, where many homes are still heated by burning oil.) Electric heaters are 100% efficient because all of the energy that goes into the heaters is turned into heat (the intended purpose). Natural gas furnaces vary considerably in how efficiently they burn gas. The most efficient ones turn about 95% of the gas energy into heat (the intended purpose); the other 5% of the gas energy is wasted through the furnace exhaust. Comparing the numbers (100% versus 95%), one could argue that electric heaters are slightly better than even the most efficient gas furnaces. Explain what is wrong with this argument, using numbers to support your answer.

(J) **Electric power again.**

(1) Some of the electricity generated by the Electricity Generation sector was successfully distributed to users, but much was lost in the system (see question (E)). Some of the electricity that is distributed to the Residential, Commercial, Industrial, and Transportation sectors is further wasted (i.e. lost). Compute the total amount of electricity that is wasted after it is transported to these 3 sectors.

(2) Determine the total amount of electric energy that is distributed and then used, and the total amount that is wasted, for the U.S. electric power system.

(3) In your answer to part (2) you were making a certain proportionality assumption about the end-use sectors. Explain.

(K) Based on the 2010 data, if you were asked to recommend *two aspects* of the U.S. energy economy where changes might increase total efficiency most, what would they be? What sorts of changes would be necessary? Would they be matters of better technology, changes in attitudes of people, etc.?

(L) Now consider the diagram for 2016.

(1) What major changes do you see between 2010 and 2016? What accounts for those changes? (If you need to look up information to answer this, as always, document your sources!)

(2) Is the U.S. energy economy more or less efficient in 2016 than it was in 2010? Explain the criteria you are using to derive your answer.

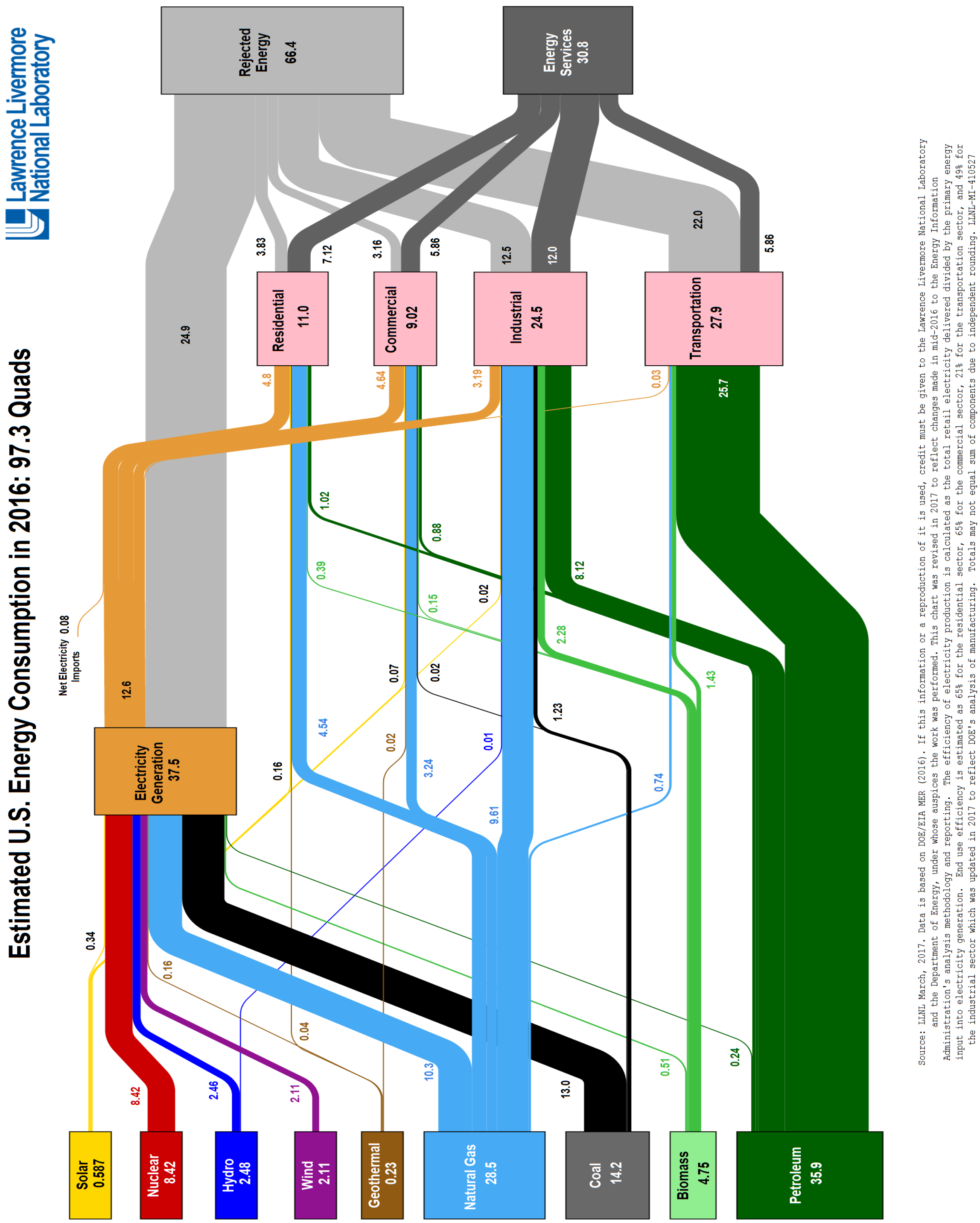


FIGURE 3.2. The U.S. energy sector in 2016.

II. Elementary Modeling

CHAPTER 4

Linear Functions as Models

Mathematicians are like Frenchmen – whatever you say to them, they translate it into their own language and it is immediately something completely different – Johann Wolfgang von Goethe

4.1. Introduction

In Table 1, we show the percentage of the U.S. new car sales accounted for by hybrid-electric vehicles (such as the Toyota Prius).[1] Hybrids incorporate internal combustion engines and electric engines powered by storage batteries. This design increases fuel efficiency (typically to levels over 40 mpg) and reduces tailpipe emissions. These vehicles became widely available in the early 2000's. Mathematically,

Table 1. Hybrid Vehicle Market Share in the U.S.

Year	2002	2003	2004	2005	2006	2007	2008
% of sales	.24	.32	.56	1.40	1.77	2.55	2.37
Year	2009	2010	2011	2012	2013	2014	2015
% of sales	2.79	2.37	2.11	3.01	3.19	2.75	2.21

we can consider the hybrid vehicle market share as a *function* of the years in this range. This means that for each year shown, the market share of hybrid vehicles during that year is a well-defined number depending on the year. The value 3.19 for the year 2013, for instance, means that 3.19% of all new vehicle sales in the U.S. were hybrids that year.

This information can be represented graphically in several ways. In Figure 4.1, we show a point plot with the year on the horizontal axis and the market share on the vertical axis. Another way information such as this might be presented is as a "bar chart," as in Figure 4.2. There we include one bar for each year, whose height gives the market share in that year. Both of these plot styles are designed to let us visualize the change in the hybrid market share over time. And they both show a rather irregular up-and-down pattern with peaks in 2009 and 2013, followed by declines. Evidently, the appeal of hybrid vehicles is different in years when gasoline prices are trending upward; one of the attractive features of hybrids is their greater fuel efficiency. But this also comes at a higher cost in the sticker price

[1] From Alternative Fuels and Advanced Vehicle Data Center, U.S. Department of Energy, `https://www.afdc.energy.gov/data/10301`, accessed June 26, 2017.

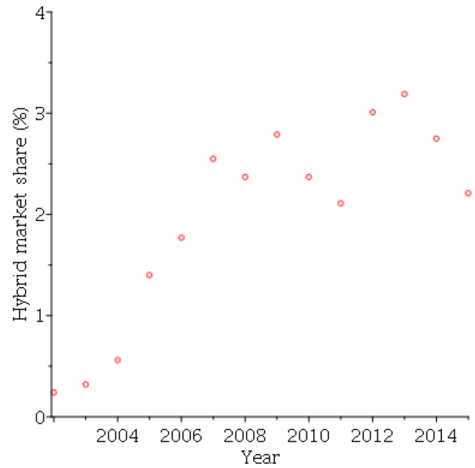


FIGURE 4.1. Point Plot of Hybrid Car Market Share Data

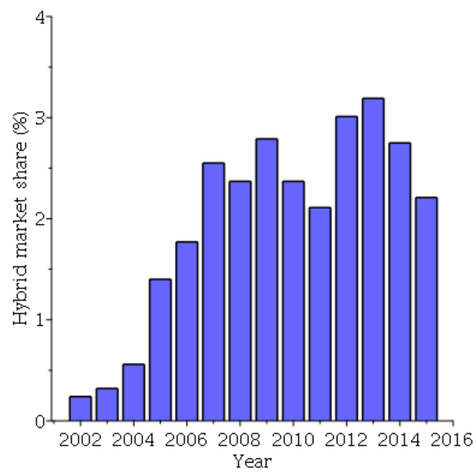


FIGURE 4.2. Bar Chart of Hybrid Car Market Share Data

of hybrid vehicles. As a result, people considering purchasing a hybrid often take their expected fuel consumption into account in making the decision and hybrids tend to be less attractive when fuel prices are low.

Nevertheless, we can see what is apparently an upward trend in the hybrid market share over time. In this chapter we will focus on modeling situations like this using what are called *linear* models, where the graph of the corresponding mathematical function would be a *straight line*. We will also see a way, called *linear regression*, to estimate the straight line that fits a data set like our hybrid market share data as closely as possible.

4.2. Functions

We will begin by discussing the general mathematical concept of a *function*. This idea may be familiar if you have taken a precalculus or calculus course, but we will review the basics and use that language here as well. A function can be thought of as a rule, or a process, or perhaps even as a sort of machine, that assigns to each element in some set (called the *domain*) a *unique* element of some other set (called the *range*). We often think of the domain D as the set of *input values* for the function; the corresponding *output values* make up the range. The most important part of this definition is that there is *exactly one corresponding output for each possible input* – there is no ambiguity or multiplicity involved in the process of going from the input to the output.

For us, the domain and range will always be sets of numbers. Mathematicians often give functions single-letter names such as f and write $f : D \to R$ to mean that D is the domain of the function and R is the range. If x is an element of the domain D, then the notation $f(x)$ represents the element of R that the function produces as output when the input value is x.

Example 4.1. Probably the most common way to specify a function is to give a *formula* for producing $f(x)$ from x.

(a) For instance the formula

$$f(x) = 17.9 \cdot x + 23.7$$

defines a function with domain the set of all real numbers because we can compute $f(x)$ for any real number by taking x, multiplying it by 17.9, then adding 23.7 to the result. For instance

$$f(1.1) = 17.9 \cdot 1.1 + 23.7 = 43.39.$$

Similarly

$$f(-2) = 17.9 \cdot (-2) + 23.7 = -12.1.$$

This function is especially simple because we can substitute any real value for x in the formula.

(b) This may not be true for other more complicated function such as the one defined by

$$g(x) = \frac{1}{x^2 - 16}.$$

Here notice that $g(4)$ and $g(-4)$ are not defined because substituting $x = 4$ or $x = -4$ makes the denominator in the formula equal to zero; division by zero is not defined. In cases like this, the usual "rule of thumb" is to take the domain to be the set of all real x that can be substituted into the formula to yield a well-defined value (unless some other domain is explicitly specified).

(c) The base-a logarithms that we discussed in Chapter 1 are values of function $f(x) = \log_a(x)$ with domain equal to the set of all strictly positive real numbers. The range of the base-a logarithm is the set of all real numbers. These statements reflect the facts that the logarithm of a negative number is not defined in the real numbers, but the value of a logarithm can be negative if the input is < 1 and $a > 1$. $\triangle$

Other ways of specifying functions will also be important for us in this course.

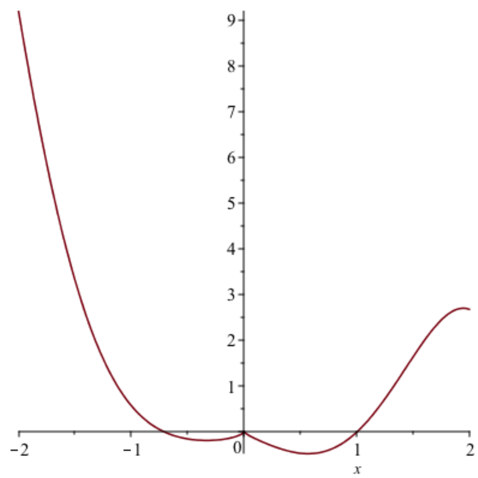


FIGURE 4.3. Function Defined by a Graph

EXAMPLE 4.2. The hybrid vehicle market share values in Table 1 in the Introduction to this chapter can be thought of as the definition of function with domain

$$D = \{2002, 2003, \ldots, 2015\}$$

and range given by the set of market share values. The other tables of values we have considered in previous chapters could be thought of in the same way. Especially when a table of values represents the data collected in a scientific experiment or measurement process, we may want to think of those values as a sort of *incomplete description* of a function that might be defined for other input values as well. For example, the tabulated data in Tables 2 and 3 from our discussion of the Weber-Fechner Law in Chapter 1 show subsets of the ranges of the $\log_{10}$ function, but that function is defined for all strictly positive x, not just for the input values given in those tables. △

EXAMPLE 4.3. A function could also be defined by giving a graph in the xy-plane where each vertical line $x = c$ meets the graph *at most once*, as in Figure 4.3. If (c, d) is the point on the graph with $x = c$, then we would define $f(c) = d$. For instance, from this plot we can see $f(0) = 0$ and $f(2) \doteq 2.7$. We say such a graph "passes the vertical line test." Even if we do not have an explicit formula for such a function, the graph is a complete specification. △

4.3. Linear Functions

In this chapter, we focus on one particular class of functions–the ones known as *linear functions*–and their applications. There are several ways to say which functions are linear and to understand what makes a function linear.

First, linear functions are exactly the functions whose formulas can be rearranged to the form

$$\boxed{f(x) = mx + b} \tag{4.1}$$

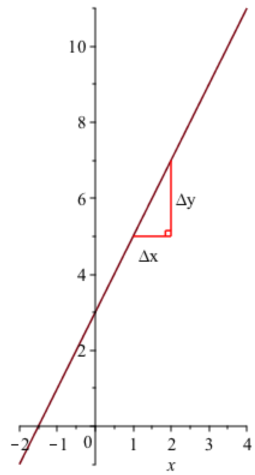


FIGURE 4.4. Slope of a Line

where m, b are constants. As you probably recognize, the graph

$$y = f(x) = mx + b$$

is just a *straight line* and that gives a second characterization of the linear functions in terms of their graphs. The constants m, b have a nice interpretation in terms of that line. The constant m is called the *slope* of the line and it has the following interpretation. Suppose we start from $x = a$ and the corresponding point $(a, f(a)) = (a, ma + b)$ on the graph of the linear function. If we change the x-coordinate to $a + \Delta x$, then the y-coordinate changes to

$$f(a + \Delta x) = m(a + \Delta x) + b = ma + m\Delta x + b.$$

Hence the corresponding change in the y-coordinate is just

$$\Delta y = f(a + \Delta x) - f(a) = ma + m\Delta x + b - (ma + b) = m\Delta x.$$

Assuming $\Delta x \neq 0$,[2] we can divide through to yield

$$\frac{\Delta y}{\Delta x} = m. \tag{4.2}$$

Notice that this equation is true for all a and all Δx. In other words, according to the language introduced in Chapter 2, for all a and all Δx, the ratios $\frac{\Delta y}{\Delta x}$ are always the same, equal to the slope m. Another way to say this is that Δy and Δx are always *proportional*, with ratio equal to m. This gives an important additional way to understand the property of linearity, which is illustrated by the diagram in Figure 4.4. *If we take any two points on the graph of a linear function and draw in a small right triangle with legs parallel to the coordinate axes and hypotenuse along the graph, as in Figure 4.4, then ratio of the legs (which is the tangent of the angle between the horizontal leg and the line) is always the same, and equal to the constant m in the equation of the line as in* (4.1). A slope value also comes with

[2]If $\Delta x = 0$ for any two distinct points on the line, then the line is vertical, and the equation has a different form, namely $x = c$ for some constant c. We do not have the graph of a function in that case because a vertical line does not pass the vertical line test.

associated units, as is true for any ratio. These are $\frac{\text{units of } y}{\text{units of } x}$. There is also an easy interpretation for the constant b in (4.1). The value $b = f(0) = a \cdot 0 + b$. So the point $(0, b)$ on the y-axis is also on the line defined by $y = mx + b$. This point is called the *y-axis intercept.* Hence b is sometimes also called the *intercept* and (4.1) is called the *slope-intercept form* of the equation of a line. Because the b is a value of y, its units are the same as those of y.

Example 4.4. Hence, for instance, the line defined by $f(x) = -3x - 1$ has slope $m = -3$ and y-axis intercept $(0, -1)$. The negative value of the slope means that the line is "sloping down" left to right – in other words, the y-coordinates of points on the line are decreasing as the x-coordinates increase. $\triangle$

An alternative way to specify a linear function is to give the slope m and one point (x_0, y_0) on the line (i.e. the graph of the function). If (x, y) is any other point on the line, then using those two points we have

$$\Delta x = x - x_0 \quad \text{and} \quad \Delta y = y - y_0$$

so from (4.2), we obtain

$$\frac{y - y_0}{x - x_0} = m,$$

and hence

$$\boxed{y - y_0 = m(x - x_0).} \tag{4.3}$$

This is called the *point-slope form* of the equation of the line. Note that this form can be rearranged to give

$$y = mx + (y_0 - mx_0)$$

by distributing the product of m and $x - x_0$ and adding y_0 to both sides. So we have a linear function $f(x) = mx + (y_0 - mx_0)$ as in (4.1), where

$$b = y_0 - mx_0.$$

There is no need to memorize this formula for the intercept; you can always rearrange an equation given in point-slope form to slope-intercept form as above if you need to find the intercept.

Example 4.5. The line with slope $m = 4.2$ containing the point $(1.1, 2.8)$ has equation

$$y - 2.8 = 4.2(x - 1.1) \quad \text{or} \quad y = 4.2x - 1.82. \quad \triangle$$

It is a basic fact from geometry that there is exactly one line passing through any two distinct points in the plane. Given the two points, (x_1, y_1) and (x_2, y_2), if we need to determine the line, the fastest way is to compute the slope using (4.2) first, and then use (4.3). The known point (x_0, y_0) can be either of the two given points; the resulting equation will be the same after simplification.

Example 4.6. Say we have the two points $(x_1, y_1) = (2.2, 8.7)$ and $(x_2, y_2) = (4.3, 5.0)$. We ask, what is the equation of the line through those points? The slope must be

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{5.0 - 8.7}{4.3 - 2.2} \doteq -1.76.$$

So, using $(2.2, 8.7)$ as the known point in (4.3), the line is (approximately)

$$y - 8.7 = -1.76(x - 2.2) \quad \text{or} \quad y = -1.76x + 12.57.$$

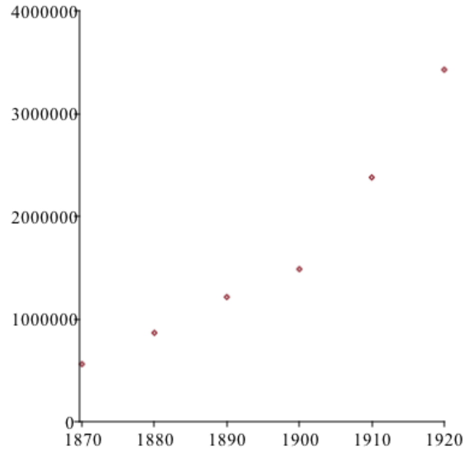


FIGURE 4.5. California Population Data

Note that if we substitute $x = 4.3$ and compute y we find

$$(-1.76)(4.3) + 12.57 = 5.002,$$

which is not exactly 5.0 as we expect for $x = 4.3$. The reason for this is that our slope $\frac{-3.7}{2.1} = -\frac{37}{21}$ is a rational number with an infinite repeating decimal expansion. Therefore the rounded value $m \doteq -1.76$ is only approximate. △

To conclude this section, we will consider how to tell whether a table of values comes from a linear function and determine the equation if it does. This basically uses all the ideas we have introduced so the idea should be clear if we just consider the following example.

EXAMPLE 4.7. The population of the state of California, measured at 10-year intervals in the period between 1870 and 1920 is given Table 2.[3] We ask whether

TABLE 2. Population of California 1870 to 1920

Year	1870	1880	1890	1900	1910	1920
Population	560,247	864,694	1,213,398	1,485,053	2,377,549	3,426,861

these are values of some linear function of time in years. According to (4.2), that would be true exactly when the slope between any two of the points was the same. However, using the populations for 1870 and 1880 gives

$$\frac{864,694 - 560,247}{1880 - 1870} = 30,444.7 \text{ people/year},$$

while the same computation for the populations in 1880 and 1890 gives

$$\frac{1,213,398 - 864,694}{1890 - 1880} = 34,870.4 \text{ people/year}.$$

[3]According to U.S. Census data.

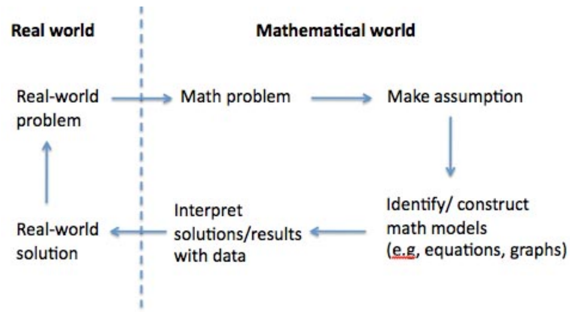


FIGURE 4.6. The Mathematical Modeling Process

This alone is enough to show that the tabulated values *do not* come from any single linear function.

Nevertheless if we make a point plot of these population values as in Figure 4.5 then we see some interesting things. First notice that the points here are not actually all that far from lying on a straight line. We could ask – Is there a line that "fits" the data as well as possible. Looking more deeply, we can also see that there was apparently also some sort of "change point" in 1900. The first four points from the census values for 1870, 1880, 1890 and 1900 are actually even closer to lying on one line than the data points as a whole. The same is true for the final three points from the census values for 1900, 1910, and 1920. We might try to "fit" two lines separately and compare their slopes to understand the apparent difference between the first part of the time period and the second part. We will return to this example later in this chapter. △

4.4. Mathematical Models

Many of the mathematical topics we will study this year will be applied to construct *mathematical models* of various real-world and environmental situations. The idea of mathematical modeling is shown in the diagram given in Figure 4.6.[4] This sort of mathematical work is (almost) always motivated by the desire to understand and/or make predictions about the behavior of some real-world system.[5]

The modeling process can be viewed as a sort of "loop" or iteration as shown by the arrows in Figure 4.6. The loop should ideally start with a deep understanding of the real-world problem, and for this reason mathematicians doing this sort of work often collaborate with biologists, demographers, climate scientists, and other experts in the applications areas they want to study. Moving into the mathematical

[4]Source: `http://mathforteaching.com`, downloaded on June 27, 2017.

[5]Some mathematicians also study mathematics "purely" for its own sake, for the beauty they see in its structures, or for other reasons, without any applications in mind. One of the surprising things we can see from the history of the subject is that even though many mathematical techniques were originally developed without any applications in mind, they have later turned out to be useful for applications even so. A famous essay by the physicist Eugene Wigner called *The unreasonable effectiveness of mathematics in the natural sciences* discusses some of this.

world, the underlying real-world problem is often reformulated and abstracted as a mathematical question. This often involves making simplifying assumptions or leaving out or ignoring aspects of the real world system that are thought to be irrelevant (or at least less important) for the problem at hand.

A mathematical model can be an equation, or a graph, or some other mathematical structure that captures some of the features of the real-world system under study. The mathematical structures make it possible to compute solutions of equations or produce other mathematical results. Is that the end of the story? *Definitely not.* These results from the realm of mathematics must be *tested against*, or compared with, data generated from the real-world system.

- If there *is* sufficiently close agreement or the mathematical results yield good insights about that system, then the model has produced useful information. Even in this case, though, we must always be ready to acknowledge that there may be other important processes in play that our models are not not taking into account. We may consider refining and/or completely scrapping our models because of that.
- If there is *not* the sort of agreement we are seeking, and in many cases, even if there is, the modeling loop goes through additional iterations to refine the models and improve their predictive power. In particular, it might be the case that some features ignored previously are seen to be essential to the development of a realistic model and useful predictions.

Unlike the case in pure mathematics where solving an equation, or developing a proof of a statement can be the end of the story, mathematical modeling can be a much more incremental process–much more like the sort of experimental work that goes on in many scientific disciplines, in fact.

Some of the simplest mathematical models that are used in practice make use of the linear functions we introduced earlier. A linear model says that one quantity, the *dependent variable* corresponding to the y, depends on another quantity, the *independent variable* corresponding to x, in a linear fashion as in $y = mx + b$, at least *approximately.*

EXAMPLE 4.8. We continue the discussion of the California population data from Example 2 above. Recall that we said that, in the point plot from Figure 4.5, the population values for 1870, 1880, 1890, and 1900 were visually very close to being collinear. We will see in the next section how to derive an equation of the line that "best fits" this data, and using the techniques to be presented there we will be able to derive the equation

$$\text{Population} \doteq (31231.22) \cdot (\text{Year}) - 5.784 \times 10^7 \tag{4.4}$$

for this best fit line. The closeness of the line defined by this equation and our data points can be seen if we plot both together on the same axes as in Figure 4.7.

Our main point is that we can view this linear function as a *mathematical model* for the California population, at least during the period between 1870 and 1900. A portion of the meaning of this would be that the population was growing at a constant rate of about $31,231$ people per year, from the numerical value of the slope coefficient in the linear function from (4.4). Having this model would, for instance, allow us to generate estimates for the population in years where there

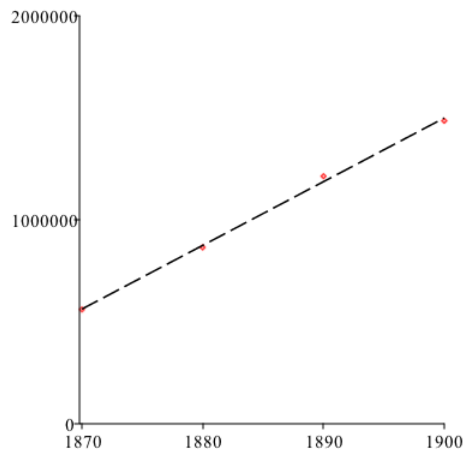


FIGURE 4.7. A First "Best Fit" Line

was no Federal census. For instance with Year = 1882, we obtain

$$\text{Population} \doteq (31,231.22)(1882) - 5.784 \times 10^7 \doteq 937,154.$$

Looking back at Table 2, this is clearly reasonable as an estimate for the population that year.

On the other hand, as we noted before, something seemingly changed in the way the California population was growing around 1900. The value obtained from the linear function in (4.4) for Year = 1920,

$$(31,231.22) \cdot (1920) - 5.784 \times 10^7 \doteq 2,123,941$$

is much smaller than the actual population that year, namely 3,426,861. Similarly, extrapolating back to values like Year = 0 yields *negative* population estimates(!) Hence predictions from a mathematical model must be used with care. Here the operative restriction to keep in mind is that the formula in (4.4) was derived using only the actual population values in the period from 1870 to 1900 so for it to produce reliable estimates, it should essentially only be applied for Year values in or close to that time range. $\triangle$

4.5. Linear Regression

In this section we will present an overview of the technique of *least squares regression* which is the standard method used to compute the equation of a line, or *linear model*, best fitting a collection of data points as in (4.4) in Example 4.8. The computations involved are somewhat tedious to perform by hand, so after giving a simple example, we will also discuss how they can be done in an Excel spreadsheet. It is good to understand what is really going on "under the hood" there, though, so we will not treat least squares *only* as a "black box."

The basic problem is this: Given a collection of points (x_i, y_i), for $i = 1, \dots, N$, find the equation of the line that "comes the closest" to passing through the points in a suitable sense. For the moment, we consider a general line $y = mx + b$, where we think of the slope m and the intercept b as variables whose values determine which line we are talking about. In most cases we think of the x_i as *known exactly* since the independent variable x is something we, or whoever collected the data we are using, had control over in the data collection. For instance, if x represented a quantity like a temperature and y represented the proportion of insect larvae that survive to adulthood as a function of the temperature, then measurements were probably made at several different temperatures x_i and the proportions y_i were determined for each one. Similarly, in the California population data set, there is no uncertainly in the x_i – they are just the years when the Federal census is taken. On the other hand, the y_i typically do involve some randomness or uncertainty such as measurement error.

Hence it makes sense to compare each y_i with the y-value predicted by the linear function $mx + b$ with the input value $x = x_i$. The comparison is made by means of the difference or *deviation* of y_i from the value $mx_i + b$ predicted by the model:

$$y_i - (mx_i + b).$$

In graphical terms, we can think of this as the vertical distance between the point (x_i, y_i) and the point on the line $y = mx + b$ with $x = x_i$. Least squares regression determines the line–equivalently the pair of values of m and b–that *minimizes the sum of the squares of the deviations*:

$$S = \sum_{i=1}^{N} (y_i - (mx_i + b))^2 = (y_1 - (mx_1 + b))^2 + \cdots + (y_N - (mx_N + b))^2. \tag{4.5}$$

Squaring the deviations makes small deviations[6] "count less" and large ones[7] "count more." Minimizing S–that is, essentially, varying m, b to make S as small as possible–means we are eliminating large deviations as much as possible. Squaring also means that underestimates, where the deviation is positive, are *not canceled out by overestimates*, where the deviation is negative. This is reasonable if we are looking for a line that is as close as possible to passing through all of the points (x_i, y_i) but we are willing to allow the line to pass over or under any one of those points.

The one technical point that we will not try to explain here is the following. Techniques from calculus (or linear algebra) show that the m, b values that achieve the minimum for S are always given by the solutions of the two so-called *normal equations*:

$$\begin{aligned} \left(\sum_{i=1}^{N} x_i^2\right) \cdot m + \left(\sum_{i=1}^{N} x_i\right) \cdot b &= \sum_{i=1}^{N} x_i y_i \\ \left(\sum_{i=1}^{N} x_i\right) \cdot m + N \cdot b &= \sum_{i=1}^{N} y_i. \end{aligned} \tag{4.6}$$

This gives a system of two simultaneous linear equations for m and b, and we can solve them with techniques from high school algebra to find the m and b values in the equation of the least squares regression line.

[6] i.e. those less than 1 in absolute value.

[7] i.e. those greater than 1 in absolute value.

Example 4.9. In order to be able to say one really understands least squares regression, everyone should do one of these calculations at least once *by hand*. Let us consider the simple "made-up" data set from Table 3. The idea is that we can

Table 3. A hypothetical data set

i	1	2	3	4	5
x_i	1.2	1.9	2.1	2.5	3.7
y_i	5.4	6.6	6.4	7.1	6.9

simply write down the normal equations from (4.6) without going through the whole computation of the total squared deviation. We have 5 data points, so $N = 5$. Then

$$\sum_{i=1}^{N} x_i^2 = (1.2)^2 + (1.9)^2 + (2.1)^2 + (2.5)^2 + (3.7)^2 = 29.40$$

$$\sum_{i=1}^{N} x_i y_i = (1.2)(5.4) + (1.9)(6.6) + (2.1)(6.4) + (2.5)(7.1) + (3.7)(6.9) = 75.74$$

$$\sum_{i=1}^{N} x_i = 1.2 + 1.9 + 2.1 + 2.5 + 3.7 = 11.4$$

$$\sum_{i=1}^{N} y_i = 5.4 + 6.6 + 6.4 + 7.1 + 6.9 = 32.4.$$

Hence the normal equations are

$$\begin{aligned} 29.4m + 11.4b &= 75.74 \\ 11.4m + 5b &= 32.4. \end{aligned}$$

By elimination or Cramer's Rule, the solution is

$$\hat{m} \doteq .548 \quad \text{and} \quad \hat{b} \doteq 5.23.$$

We will consistently write $\hat{m}$ and $\hat{b}$ for the solution of the normal equations in a linear regression.[8] The data points and the (approximate) regression line $y = .548x + 5.23$ are plotted in Figure 4.8. The relatively close agreement between the dashed black line and the red data points indicates that this linear model fits the data quite well. $\triangle$

We now discuss a basic way to generate point plots, compute equations of regression lines, and plot them with the point plots in Excel. (There are several other ways to do this as well; this is the simplest method.)

To create a point plot of what Excel wants to call a *bivariate data set* (that is a collection of (x_i, y_i) points for $i = 1, \ldots, N$), you will follow these steps:

(1) Enter the x_i and y_i values into the spreadsheet in two consecutive columns. (To help you understand what you did if you come back to the spreadsheet later, it is often helpful to enter text headings in the cells at the top of the columns, but that is not necessary.)

[8] This notation is borrowed from the way statisticians indicate an *estimator* for a parameter in a statistical model – here estimators for the slope and the intercept in the linear model.

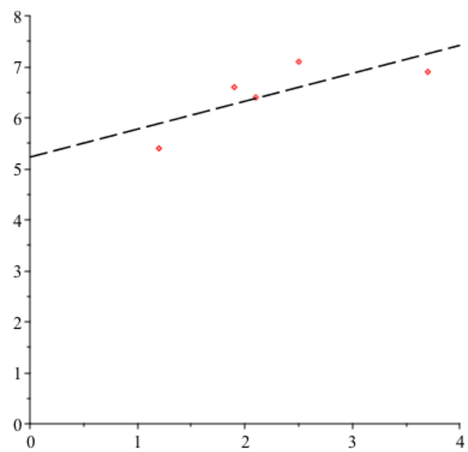


FIGURE 4.8. A "Best Fit" Regression Line

(2) Highlight the range of cells containing the data. (On a PC you do this by holding down the left "mouse" button and dragging the cursor. On a Mac, use one finger to highlight and drag with the other.)
(3) With the data highlighted, press the Insert tab of the Excel window, and choose the option Scatter from the Plots group.
(4) You should see a bare-bones version of the plot generated at this point.

You will almost always want to edit your plot to add axis labels, a chart title, regression lines (examples of what are called *trend lines* in Excel), equation(s), etc. To do this you will use various options in the Layout tab of the Chart Tools (PC) or Chart Design (Mac) group.

(5) With the Layout tab of the Chart Tools group highlighted on a PC or the Add Chart Element tab on Mac, you should see options labeled Chart Title, Axis Titles, Legend, etc. Each of those buttons produces a menu that you use to add or remove features of the chart. The title options, for instance, add text boxes overlaying the graph that you type in to add the title you want.
(6) The Trendline menu contains options for linear regression and other sorts of calculations, some of which we will use later. The Linear Trendline button just adds the trendline, though. If you want to be able to generate the equation of the line overlaid on the plot, go to the bottom option in the menu (Other Trendline Options), select the trend/regression type you want, and check the box that says Display Equation on Chart (this is down at the bottom of the list and you may need to scroll down to find it).

Figure 4.9 shows a screen shot of the Mac version of Excel 2016 showing results of commands like those given above on the data set entered in cells `B2` to `C7`. The chart title *Regression Example* was generated automatically by the column heading in column `C` (the contents of cell `C2`). The x-values and y-values labels on the axes in the plot were added manually. The equation $y = 2.8226x - 6.302$ is in a "text

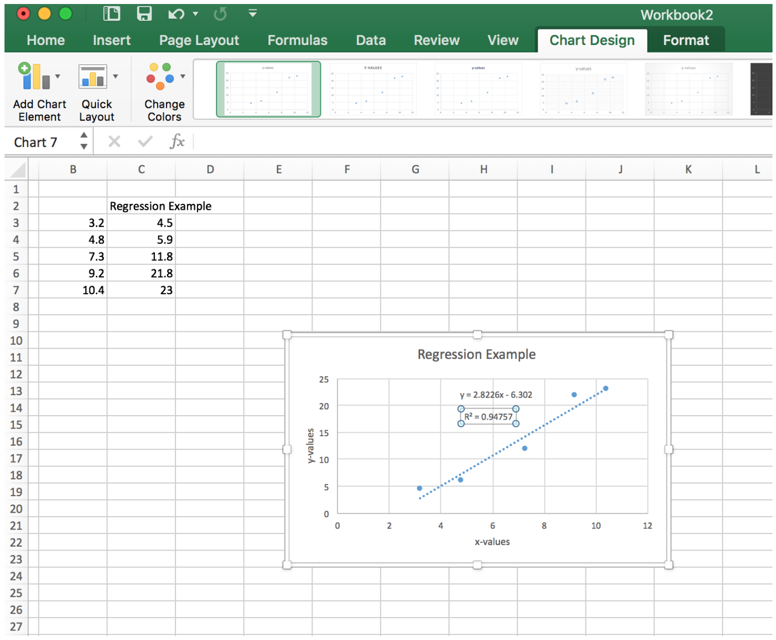


FIGURE 4.9. A "Best Fit" Regression Line in Excel.

box" element. This may come out covering part of the graph but you can move it. The same is true of the equation $R^2 = .94757$. The R^2 statistic is a measure of how close the regression line comes to passing through the data points. The closer that is to 1, the better the "fit," so this is quite good. We will discuss the R^2 statistic, its precise meaning, and how it is computed in more detail in Chapter 9.

When evaluating whether a linear model is a good fit for a data set, experienced mathematical modelers and statisticians usually do an additional step – investigation of the so-called *residuals* for the regression. The residuals are the differences

$$r_i = y_i - (\hat{m}x_i + \hat{b}),$$

where $\hat{m}, \hat{b}$ are the solutions of the normal equations from (4.6). The residual r_i is a measure of how far "off" the value for y with $x = x_i$ predicted by the linear model is from the measured value y_i from the data set. The *residual plot* is the point plot (Excel scatter plot) for the data set consisting of the points (x_i, r_i). If the scatter plot shows more or less random-looking up and down variations from 0, then there is no reason to be concerned. On the other hand, if the residuals show some systematic pattern–for instance,

- if they are all negative for the first half of the data, then all positive for the second half, or more generally
- if the scatter plot has long intervals of consistent curvature creating consistent sign patterns in the residuals,

then the linear model might be missing something and use of another type of model might be suggested.

For example, in Figure 4.9, even without computing the numerical values of the residuals, we can see that r_1 will be slightly positive (the first point has y_1 greater than $\hat{m}x_1 + \hat{b}$ since the point (x_1, y_1) lies above the regression line), then r_2 is slightly negative, r_3 is slightly more negative than r_2, r_4 is positive and finally r_5 is quite close to zero since that point comes close to lying on the regression line (it's actually slightly postive, but quite a bit less positive than r_1 or r_4. This is more or less the sort of "random up and down" pattern that indicates a linear model is reasonable. A few words of caution are probably in order here, though. Our comments above apply most directly when N is large. $N = 5$ is probably too small to talk about any sort of patterns. Moreover, when examining residuals for a linear model, it is quite easy to start seeing patterns that are probably only chance variations–our visual perception systems have evolved to work that way because there was probably some advantage to perceiving such patterns even when they are not really there.[9]

4.6. Chapter Project

Background Information. The Mauna Loa Observatory (located at an elevation of about 3400 meters on the Mauna Loa volcanic mountain on the "big island" of Hawaii) is a research station maintained and staffed by the National Oceanic and Atmospheric Administration (NOAA), the major agency of the US government that studies weather and climate phenomena. NOAA maintains a web site, ! for instance, with the National Weather Service day-to-day forecasts and severe weather warnings that form the basis for most local weather reporting in the media. Among the data collected regularly at Mauna Loa are measurements of atmospheric concentrations of a host of trace gasses, including carbon dioxide, or CO2. The data set of those measurements goes back to 1958 and is one of the most complete records of the recent evolution of this aspect of the Earth's atmosphere. In this project we will apply the modeling techniques introduced in this chapter to try to understand what this data set is saying about changes in atmospheric CO2 over time.

Important Note. This is a well-known data set and you can find all sorts of discussions of various aspects of it on the web, if you look. I am going to ask all of you *not to look* until after you have worked through at least questions A through D below, though. The idea is for you to approach this entirely "fresh" and make your own observations and analysis and draw your own conclusions.

Getting Started. The data we will be looking at is contained in a (large) Excel spreadsheet called `MaunaLoaCO2Data.xls` that you will download from the course homepage. Begin by getting the spreadsheet and opening it in Excel. Note the layout:

- Column `A` gives the year the measurement was taken
- Column `B` gives the month (1 = January, through 12 = December)

[9]Think of the possible value of a *flight response* generated by perceiving a predator in the trees in the distance, even when there is really nothing there!

- Column `C` gives a decimal equivalent of the middle of the month, so for instance January 1958 is given as 1958.04, since 1 month = 1/12 year = .08 year (roughly), and .04 year is about 15 days.
- Column `D` gives the average CO2 level observed at Mauna Loa that month in units of parts per million
- In Column `D`, if you look closely, you will see that a few of the entries near the start are -99.99. What do you suppose that means?
- In Column `E`, you will notice that most of the entries are the same as the corresponding entries in Column `D`, but the -99.99 entries have been replaced by other values. These are "interpolated" (estimated) values based on the trends from the nearby months. We will use Column `E` for all our values so that the -99.99's are not included.

Questions. The first thing you will notice if you look at the CO2 levels is that there is *a lot* of up-and-down variation. Is it completely random, though? And is there an underlying trend?

(A) To start to answer this question:

(1) Create scatter plots of the CO2 monthly averages for the calendar years 1965, 1975, 1985 (individually), versus the decimal year from Column `C`. This will require picking out the correct range of rows in Columns C and E for each of these years, and you may want to copy those values to other cells to create the scatter plots.

(2) Looking at these scatter plots, what do you notice about the way CO2 levels vary over these years? Describe what happens over the course of a typical year, and hypothesize a reason why the annual pattern works this way. Note: Mauna Loa is in the Northern Hemisphere and typical mixing patterns in the atmosphere mean that most of the air that passes over this location has come from other areas in the Northern Hemisphere. What happens through the months of May, June, July, August in the Northern Hemisphere, and how might that affect atmospheric CO2 levels?

(3) *Extra Credit*: How might you model the yearly variation of the CO2 readings? Suggest mathematical function(s) that might be useful and how you might apply them.

(B) Condensing the Data to a More Manageable Form. Our goal is to model how atmospheric CO2 levels have been changing over this period (but on the year-to-year level, not on the much more variable month-to-month level). This will be much more manageable if we identify some way to compute a "summary value" for each year to use as the representative CO2 level for that year.

(1) Identify (at least) three different ways that might be used to produce that sort of "summary value" and describe why they would be suitable.

(2) Choose one of your proposed ways to do this and give a reason for why you think that will be a reasonable way to "condense" the data for each year.

(3) Create new columns in your spreadsheet giving the number of years since 1959, and your summary CO2 value for the year. Since we don't have complete values for the years 1958 or 2017, just use the years 1959 to 2016 (58 years in all).

(C) "Let the modeling begin!"

(1) Using Excel, fit a linear model to your "condensed" data set and record your results. Give the equation of the regression line as a function of the years since 1959 (that is, $x = 0$ correspond to 1959, $x = 1$ corresponds to 1960, and so forth). Also give the R^2 value reported by Excel as a measure of goodness of the fit, and discuss the residuals for the linear model (in particular, is there a consistent pattern there)?

(2) What does your model predict concerning the CO2 level in 2020? (This is slightly outside the interval 1959 to 2016 of course, but not too far outside. So extrapolation from the linear model is at least a possibility!)

(D) Atmospheric CO2 levels are of concern, of course, because of the "greenhouse gas" properties of this compound–the way atmospheric CO2 can trap energy from reflected solar radiation and increase temperatures near the surface of the Earth. Some greenhouse effect is necessary for life on Earth, of course (our water-based form of life could not exist at the temperatures that would prevail with no greenhouse effect at all because all water would be frozen as ice). But have there been times in the past when CO2 concentrations were significantly higher than they are now? What were the Earth's climate and sea levels like then? (This may require some research–be sure to give the sources you used to compile your information.)

Assignment. Submit your edited Excel spreadsheet with the data and write up answers to the questions above in a separate document.

Chapter Exercises

(1) Consider the function defined by $f(x) = \sqrt{x-6}$. What is the domain of f according to the "rule of thumb" discussed in the text.

(2) Same question as in (1) for $g(x) = \frac{x}{x^2-5x+6}$.

(3) Does the equation $x^2 + y^2 = 9$ define y as a function of x? (Hint: What curve in the plane is defined by that equation?) If yes, plot that function. If not, explain why not and say how you could find subsets of the curve defined by this equation that are graphs of functions of x and give their domains.

(4) We discussed an example of a family of circles with areas in an *arithmetic* progression in Figure 1.1. Formally, an arithmetic progression is any sequence of numbers of the form

$$a, a+d, a+2\cdot d, a+3\cdot d, \ldots, a+n\cdot d$$

where a, d are fixed numbers. Show how such an arithmetic progression can be obtained as the values $f(0), f(1), \ldots, f(n)$ for some linear function. (What is the formula of that linear function?)

(5) Given that the line plotted in Figure 4.4 passes through the points $(1, 5)$ and $(2, 7)$, what is the equation of the corresponding linear function $f = mx + b$?

(6) Here are some additional problems concerning equations of lines.

(a) What is the equation of the line passing through the point $(3.2, 5.2)$ and parallel to the line given by $y = 7x + 1$? (Hint: What is true about the slopes of parallel lines?)

(b) What is the equation of the line passing through the point $(0, 1)$ and perpendicular to the line from part (a)? (Hint: The slopes of perpendicular lines are negative reciprocals of one another.)

(7) In least squares regression, we minimize the sum of the squares of the vertical distances $y_i - (mx_i + b)$ in terms of m, b to find the regression line. Another way to think about the distance from a point to a line is to consider the straight line distance from the point to the *closest point on the line*.[10] In a formula, if the point P is (x_0, y_0) and the line L is given by an equation written in the form $Ax + By + C = 0$, the distance from P to L measured this way is

$$d(P, L) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}$$

Find the distance from $P = (1, 2)$ to the line $y = 5x + 7$ measured this way.

TABLE 4. Three tabulated functions

x	1.1	2.4	3.7	5.0	6.3	7.6
$f(x)$	1.445	1.844	2.168	2.449	2.702	2.933
$g(x)$	1.525	−1.400	−4.325	−7.250	−10.175	−13.100
$h(x)$	1.111	1.257	1.423	1.611	1.823	2.063

(8) Only one of the functions tabulated in Table 4 is linear. Identify which one it is and find the equation of the linear function.

(9) Refer to Exercise 16 from Chapter 1. Construct a linear model for the area in the U.S. covered by kudzu as a function of time in years since 1876. What are the units of the slope and what does the slope represent?

(10) Compute the residuals for the linear model from Example 3 and discuss them in relation to Figure 4.8.

(11) Use Excel to find the least squares regression line for the hybrid vehicle market share data from Table 1. What is the slope of the regression line, what are its units, and what does that slope represent?

(12) In this problem we continue the study of the California population data from Table 2.

(a) Using Excel (or doing the calculations by hand), find the regression line for the subset of the data set corresponding to the years 1900, 1910, and

[10] That is different unless the line is horizontal!

1920. How does the slope compare to the slope found for the data from 1870 through 1900? Does this match what you thought by looking at Figure 4.5?

(b) Now find the regression line for the whole data set (the years 1870–1920). Plot this together with the point plot (or Excel scatter plot) and discuss the results.

(c) Find the residuals for this linear model and generate a point plot (scatter plot). Is there a systematic trend?

(13) By researching U.S. Federal Census data online,

(a) Find the population of your home state for the censuses in 1900 through 2010. (If your home state only joined the Union after 1900, use only the census data for the years since statehood.)

(b) Find the least squares regression line and the R^2 statistic.

(c) Compute the residuals for your linear model and generate a point plot (scatter plot). Examine them for any patterns as we discussed in the text.

(d) What is your conclusion? Does a linear model fit your state's change in population over this period (at least approximately)?

CHAPTER 5

Exponential Functions as Models

5.1. Introduction

In everyday speech, the phrase "it's growing exponentially" seems to have become quite common in the early 2000's. Most of us probably understand saying that to be equivalent to the assertion that whatever is being described is growing *really fast.* In mathematical modeling, this statement is also frequently made, but it has a much more precise meaning. Exponential models can indeed describe quantities that grow rapidly. But that is not always the case–other exponential models describe exponential decay where a quantity decreases and approaches zero. Moreover, even when we have exponential growth in the technical sense, we can also make much more precise and quantitative statements about how exactly how the growth works and changes over time.

Example 5.1. For organisms living in a habitat with unlimited resources,[1] in each time period of a fixed length, we would expect there to be numbers of births and deaths proportional to the total size of the population and this assumption is the basis for a new, non-linear of mathematical model describing how the population changes over time.

For instance, for populations of animals with life spans longer than a year, we often take the time in years; for bacteria, insects, etc. we might find a shorter time interval to be more convenient. The ratios

$$b = \frac{\text{live births per year}}{\text{total population}} \text{ and } d = \frac{\text{deaths per year}}{\text{total population}}.$$

are called the *birth and death rates.*[2] Over the short run, we usually assume these remain constant. The difference $b - d$, which reflects both additions to the population via births and subtractions from the population due to deaths, is then called the *net growth (or decay) rate* (per year).

If the population was $P(0)$ at the end of some year, then in the next year there would be $b \cdot P(0)$ births and $d \cdot P(0)$ deaths. Hence, at the end of the next year the population $P(1)$ would be given by

$$P(1) = P(0) + b \cdot P(0) - d \cdot P(0) = (1 + b - d)P(0). \tag{5.1}$$

Similarly, at the end of the second year, substituting from (5.1), the population would be

$$P(2) = (1 + b - d)P(1) = (1 + b - d) \cdot (1 + b - d)P(0) = (1 + b - d)^2 P(0).$$

[1] This is clearly not a realistic assumption over the long run, but it might be essentially the case in the short run as long as the population of the organisms is sufficiently small.

[2] Biologists and demographers sometimes normalize these differently, looking at the ratios per 1000 population.

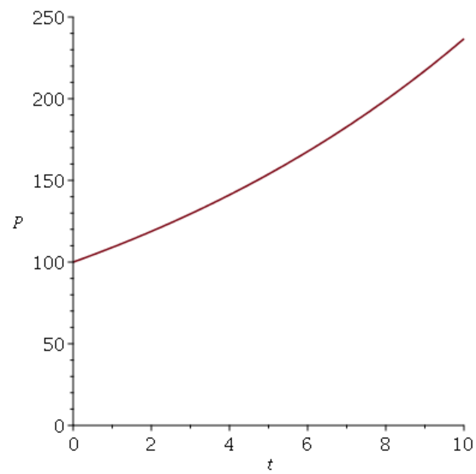


FIGURE 5.1. The Model $P = 100(1.09)^t$, $0 \le t \le 10$.

After any number $n \ge 0$ of years, in fact, the same reasoning gives

$$P(n) = (1 + b - d)^n P(0). \tag{5.2}$$

If $P(0) =$ some constant c and we extend Formula 5.2 to values of $t \ge 0$ that are not necessarily whole numbers, then we see

$$P(t) = c(1 + b - d)^t,$$

where the constant $c = P(0)$. This is a first example of an *exponential model*, so called because functions of this form are examples of exponential functions of t.

If the birth rate exceeds the death rate, that is, if $b > d$, then it is easy to see that $b - d > 0$ and $1 + b - d > 1$. For instance, if $b = .24$ (This would say 24 births per 100 population per year) and $d = 0.15$ (15 deaths per 100 population per year),[3] then $b - d = .09$, and $P(t) = c(1.09)^t$. A plot with $c = 100$ for $0 \le t \le 10$ is shown in Figure 5.1. This is the situation called *exponential growth*, and note that the population is indeed increasing more and more rapidly as t increases.

On the other hand, the death rate could exceed the birth rate in some situations, for instance if the species considered is prey for some other predator species and their reproduction is not fast enough to replace the individuals lost due to the predation. In this situation $b < d$, so $b - d < 0$ and hence $1 + b - d < 1$. For instance, if $b = .14$ and $d = 0.20$, then $b - d = -0.06$, and $P(t) = c(.94)^t$. A plot with $c = 100$ for $0 \le t \le 10$ is shown in Figure 5.2. This is the situation called *exponential decay*. The population function is decreasing and in purely mathematical terms, its values will get closer and closer to 0.

Note that the functions $P(t) = 100(1.09)^t$ and $P(t) = 100(.94)^t$ will have many values that are not whole numbers. This makes perfect sense in the "Mathematical World" in Figure 4.6. But it gives real-world predictions of the population that do not make strict sense, because a population is always a whole number of individuals and the population changes by adding or losing single individuals. The population

[3]These values would indicate a species with a relative short life span and a high turn-over in the population!

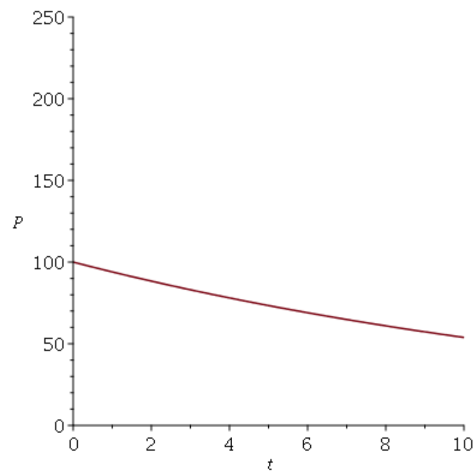


FIGURE 5.2. The Model $P = 100(.94)^t$, $0 \le t \le 10$.

model thus does not match the properties of the real-world population in every respect. Nevertheless, the agreement is close enough to give useful information in many situations. To make the transition between the model and its implications, we usually simply round the value of $P(t)$ to yield a population that is a whole number of individuals.[4] $\triangle$

In this chapter we will develop *exponential models* in general and study their properties and applications.

5.2. Exponential Functions

As is the case with linear functions, exponential functions can be defined and characterized in various ways. The most basic specifies the formulas that define these functions. Thinking in those terms, we say f is an exponential function if there is a constant $a > 0$ such that $f(x) = a^x$. The domain of one of these functions contains all real numbers. For modeling purposes we will usually generalize this slightly to include a constant c multiplying the a^x, so we will also say any function defined by an equation

$$\boxed{f(x) = c \cdot a^x} \tag{5.3}$$

is exponential. As we saw in Figures 5.1 and 5.2,

$$\boxed{a > 1 \Rightarrow \text{exponential growth}} \quad \text{and} \quad \boxed{0 < a < 1 \Rightarrow \text{exponential decay.}}$$

In the first case, the larger a is, the faster the exponential growth is. In the second case, the closer a is to zero, the faster the exponential decay is.

[4]Models of this form can also be developed for the total *biomass* of the population rather than the number of individuals. There is no restriction to whole number values necessary in that case.

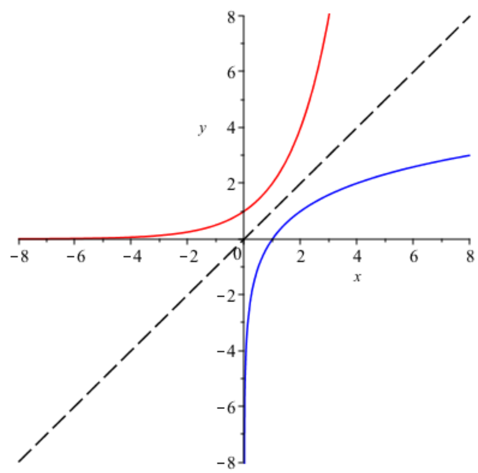


FIGURE 5.3. $y = 2^x$ in red and $y = \log_2(x)$ in blue. The line $y = x$ is shown dashed in black.

Looking back at (1.1), we can see that the a^x functions are closely related with the base-a logarithms because

$$y = a^x \Leftrightarrow x = \log_a(y).$$

In mathematical terms, this says that $f(x) = a^x$ and $g(x) = \log_a(x)$ are *inverse functions*. Each of these functions "undoes" the other and returns the input that produced a given output value from the other function:

$$f(g(x)) = x \quad \text{and} \quad g(f(x)) = x,$$

or more specifically, for all real numbers x

$$\boxed{\log_a(a^x) = x} \tag{5.4}$$

and for all positive real numbers $x > 0$,

$$\boxed{a^{\log_a(x)} = x.} \tag{5.5}$$

The pair of functions $f(x) = 2^x$ and $g(x) = \log_2(x)$ are plotted in Figure 5.3. In high school algebra or precalculus, you may have learned that graphs of inverse functions are mirror images across the line $y = x$. This is visible in Figure 5.3.

In practical terms, the fact that a^x and $\log_a(x)$ define inverse functions means that logarithms are very useful for solving equations involving exponential functions, that is, where the variable appears in an exponent.

EXAMPLE 5.2. The exponential function $f(x) = 3.4 \cdot 3^x$ grows with x and has range equal to all positive real numbers.[5] For instance, suppose we want to determine the x for which

$$10 = 3.4 \cdot 3^x. \tag{5.6}$$

[5]Technical note: This assertion actually follows from a calculus result called the Intermediate Value Theorem, but we will take it as intuitively plausible from the shape of an increasing exponential graph as in Figure 5.3.

This can be done by taking logarithms and using the statements in Proposition 1.8. One way is to simply apply a logarithm function, for instance $\log_{10}$, to both sides of the previous equation, then parts (1) and (4) of the proposition:

$$\log_{10}(10) = \log_{10}(3.4 \cdot 3^x),$$

so

$$1 = \log_{10}(3.4) + \log_{10}(3^x) = \log_{10}(3.4) + x \log_{10}(3).$$

Then we can solve algebraically for x:

$$x = \frac{1 - \log_{10}(3.4)}{\log_{10}(3)} \doteq .982.$$

Alternatively, starting again from (5.6), we could divide both sides by 3.4 and then take logarithms and use Proposition 1.8, part (4):

$$\log_{10}\left(\frac{10}{3.4}\right) = \log_{10}(3^x) = x \log_{10}(3).$$

This shows

$$x = \frac{\log_{10}\left(\frac{10}{3.4}\right)}{\log_{10}(3)} \doteq .982$$

as before. $\triangle$

Linear and Exponential Functions Compared. It follows from what we said in (4.2) that if $y = f(x) = mx + b$, is a linear function, then changing x to $x + \Delta x$ produces a change in $\Delta y = f(x + \Delta x) - f(x)$ satisfying $\Delta y = m\Delta x$. Thus,

PROPOSITION 5.3. *If $y = f(x) = mx + b$ is a linear function, no matter what x is, changing x by a fixed amount changes y by a fixed amount proportional to the change in x. Moreover, the linear functions are the only functions with this property.*

For example if $y = f(x) = -7.3x + 9$, increasing x by 1 decreases y by 7.3, since the slope is -7.3.

Exponential functions have a different, but parallel, property. Namely, if we have any $y = c \cdot a^x$ and we change x to $x + \Delta x$, now the ratio

$$\frac{ca^{x+\Delta x}}{ca^x} = \frac{ca^x \cdot a^{\Delta x}}{ca^x} = a^{\Delta x}$$

simplifies because the ca^x factors cancel. This means that:

PROPOSITION 5.4. *If $y = f(x)$ is an exponential function, then, no matter what x is, changing x by a fixed amount gives y-values that are in the same proportion.*

The change in y is not a constant in this case:

$$\Delta y = ca^{x+\Delta x} - ca^x = ca^x(a^{\Delta x} - 1);$$

it is *proportional to $y = ca^x$ (again with a proportionality constant depending on a and the change in x).* For example if $y = f(x) = 3.7 \cdot 2^x$, then increasing x by 1 *doubles* the value of y: By the rules for exponents,

$$3.7 \cdot 2^{x+1} = 3.7 \cdot 2^x \cdot 2^1 = 2 \cdot (3.7 \cdot 2^x).$$

The value for $x + 1$ is twice as large since $a = 2$ and $\Delta x = 1$.

For uses in modeling, this way of thinking about exponential functions is probably the most important one. What it says, for instance, is that an exponential function is completely determined by:

(a) *Its value for one input* value of the independent variable, and
(b) *Any one of the following*:
   (i) *A second value* at a distinct value of the independent variable, or
   (ii) *The ratio* between its value at that first value of the independent variable and the value at a second value of the independent variable, or
   (iii) *A percentage rate of growth or decrease* for a given change in the independent variable.

Here are examples illustrating how to find the formula of an exponential function from information of this kind.

EXAMPLE 5.5. (a) Suppose we know that the number of individuals in a population of some species of animals in a habitat is given by an exponential function of the independent variable t, representing time. The population is equal to 120 individuals at time $t = 0$ and that the population is declining at a rate of 3% per year. That is, at the end of each year, population is 3% smaller than it was at the start. We want to find an exponential function

$$P(t) = c \cdot a^t$$

That matches this given information. If we substitute $t = 0$, then we can determine the value of c immediately:

$$120 = c \cdot a^0 = c.$$

Then, taking t in years, the information we are given about the rate of decrease says

$$\frac{P(1) - P(0)}{P(0)} \times 100\% = -3\%.$$

Therefore

$$\frac{120 \cdot a^1 - 120}{120} = -.03 \Rightarrow a = 1 - .03 = .97.$$

Therefore, our population function is

$$P(t) = 120 \cdot (.97)^t.$$

(b) Now suppose we have a quantity Q that is increasing exponentially as a function of t in seconds and suppose we know that $Q(1) = 37.4$ and $Q(3) = 2.3 \cdot Q(1)$. We want a formula $Q(t) = c \cdot a^t$ for the corresponding exponential function again. Note that this example is slightly more complicated than (a) because we do not know the value at $t = 0$. So we cannot determine the value of c immediately as before. However, we do know

$$37.4 = Q(1) = c \cdot a^1 = c \cdot a, \text{ and}$$
$$86.02 = 2.3 \cdot 37.4 = Q(3) = c \cdot a^3.$$

and we have two simultaneous equations to solve for a, c. If we divide the second of these equations,

$$c \cdot a = 37.4$$
$$c \cdot a^3 = 86.02,$$

by the first we obtain

$$a^2 = 2.3 \Rightarrow a = \sqrt{2.3} \doteq 1.517.$$

Then the first equation can be used to solve for c:

$$c \doteq \frac{37.4}{1.517} \doteq 24.65.$$

Our exponential function is (approximately)

$$Q(t) = 24.65 \cdot (1.517)^t.$$

(c) If the two values of the exponential function are known for values of the independent variable not differing by a whole number (this was a simple feature of examples (a) and (b) above), then we might need to solve an equation like the following hypothetical relation:

$$36.5 = a^{4.23} \tag{5.7}$$

to find the value of a. This can be done by writing the exponent as $\frac{423}{100}$ and raising both sides to the $\frac{100}{423}$ power to solve for a:

$$a = (36.5)^{\frac{100}{423}} \doteq 2.34$$

Your calculator is using logarithms to do this sort of calculation; you could also take logarithms on both sides of (5.7) to obtain

$$\log_{10}(36.5) = 4.23 \cdot \log_{10}(a) \Rightarrow \log_{10}(a) = \frac{\log_{10}(36.5)}{4.23} \doteq .369,$$

so that by (5.5),

$$a = 10^{.369} \doteq 2.34.$$

A comment is probably in line here – these computations are somewhat sensitive to rounding and the final results can be different-looking depending on how many decimal digits are used. The value for a is rounded to 3 digits since the other numbers were only given to that precision. △

For future reference, we record the pattern seen in part (a) of Example 5.5 in a general form. If the values of a quantity Q are modeled by an exponential function whose value changes by $r\%$ per unit change in the independent variable t, then the model will have the form

$$\boxed{Q(t) = Q(0) \cdot \left(1 + \tfrac{r}{100}\right)^t.} \tag{5.8}$$

Note that $r > 0$ gives $a = 1 + \frac{r}{100} > 1$ and we have exponential growth. On the other hand $r < 0$ gives $a < 1$ and we have exponential decay. The percent change cannot be less than -100%, so $a > 0$.

Example 5.6. The exponential model $Q(t)$ with $Q(0) = 5.9$ and Q undergoing 2.3% growth per year is

$$Q(t) = 5.9 \cdot \left(1 + \frac{2.3}{100}\right)^t = 5.9 \cdot (1.023)^t$$

(t in years). △

5.3. Exponential Models

There are a number of situations where the use of an exponential function as a mathematical model is especially suitable. Many of these involve cases where there is a theoretical reason why exponential functions fit the data better than linear functions. For instance, it can be shown that:

Proposition 5.7. *Any function with the property that changing the independent variable x by a fixed amount gives function values y in a fixed proportion is an exponential function.*

We said above that exponential functions have this property. However, this statement is saying something different. We are claiming now that exponential functions are *the only functions* with this property.[6] Hence if we know or suspect that the property in Proposition 5.7 holds, then there is a theoretical reason to turn to an exponential model first.

The first example is exactly the situation discussed in Example 5.1 from the Introduction to this chapter. Populations of organisms in resource-rich habitats can follow exponential growth models, at least over short time-scales. Any real-world habitat is limited, though, in the number of individuals that can be supported. Exponential growth models involve population functions whose values eventually exceed any positive number. After long enough, the population size predicted by the model would be so large that the surface area of the Earth would not be big enough to hold all the organisms(!)[7] Similarly, after a long enough time, an exponential decay model will eventually reach values < 1. We cannot have .54 of an individual as a population value either! *In other words, over the long run, an exponential population model will necessarily eventually cease being realistic.*

Another situation leading to exponential models and having important environmental implications is given in the following example.

Example 5.8. Some chemical elements come in radioactive forms or *isotopes* with the same number of protons but more or fewer neutrons in each atomic nucleus than in the most common, stable forms. This includes both elements with large atomic numbers that are almost always radioactive (like Uranium and Plutonium), and other elements with smaller atomic numbers that have both stable and radioactive forms (like Carbon). A sample of a radioactive isotope will *decay* over time, whereby some of the atomic nuclei in the sample will split into nuclei of lighter elements and emit radiation of various forms. Exposure to high levels of radioactive decay can cause radiation sickness; background radioactive decay from sources in the natural environment is also always present.

The way radioactive decay works is that in a sample of a radioactive isotope, whether or not any given nucleus decays is thought to be an essentially random process. But in each time period of a given length, the *same proportion of the atoms decays* (at least approximately). Hence the number remaining (or the mass of the remaining radioactive material) satisfies Proposition 5.7 above, and this can

[6]Technical note: This is the the logical *converse* of the statement we made before and it can be shown to be true using techniques we will develop later in the course for integer values of x, and in general using calculus. We will not go into the details here.

[7]A classic *Star Trek* original series episode, with title "The Trouble With Tribbles," comes to mind. If you haven't seen it, you should :)

be modeled by a function of the form

$$Q(t) = Q(0)a^t$$

for some $0 < a < 1$. Values of a close to 1 correspond to "slow" decay, while values of a close to 0 correspond to "fast decay." The way a value for a is usually determined is often by specifying the so-called *half-life* of the element.

- *The half-life of a radioactive isotope is the time it takes for one-half of a sample to decay, or equivalently, the time it takes for the remaining portion of sample to be reduced to one-half of the original amount.*

(a) For instance the half-life of the isotope Uranium-235[8] is $7.038 \times 10^8 = 703,800,000$ years. Only slightly unrealistically, let us assume that the Earth had some amount $Q(0)$ of Uranium-235 when it first coalesced (approximately) 4.543×10^9 years ago. We can ask, how much of that original amount is left at present? First we need to determine the value for a in the exponential model $Q(t) = Q(0)a^t$ for the amount of Uranium-235 present at time t. By the definition of the half-life, we have

$$\frac{1}{2}Q(0) = Q(0)a^{7.038\times 10^8}$$

so dividing out the $Q(0)$ and taking logarithms, we can solve for a:

$$-\log_{10}(2) = (7.038 \times 10^8)\log_{10}(a),$$

hence

$$\log_{10}(a) = \frac{-\log_{10}(2)}{7.038 \times 10^8} \doteq -4.277 \times 10^{-10}.$$

Raising 10 to the power on each side of this equation,

$$a \doteq .99999999901518.$$

Then since it has been about 4.543×10^9 years since the beginning of the Earth's lifetime, we have

$$Q(4.543 \times 10^9) \doteq Q(0) \cdot (.99999999901518)^{4.543\times 10^9} \doteq .0114 \cdot Q(0).$$

In other words, about 1% of the original Uranium-235 is still left today, 4.543 billion years later.[9] This is *extremely* slow decay and it illustrates one of the issues concerning long-half-life radioactive elements in spent nuclear fuels, waste products of other processes, etc. Measurable amounts of them tend to stay around an extremely long time! This is one of the main technical issues in handling nuclear waste products and one of the reasons why nuclear power generation is controversial in environmental discussions.

(b) Here is another related example. The half-life of radioactive Cobalt-60 is about 5.27 years. We ask – how long will it take for a sample of Cobalt-60 to decay to 5% of the original amount? Our exponential model is $Q(t) = Q(0)a^t$. The information about the half-life says

$$Q(5.27) = Q(0)a^{5.27} = .5 \cdot Q(0).$$

[8]This isotope is used in fuel for nuclear power plants and in nuclear weapons because it is "fissile"– that is, under the right conditions, a large enough sample will produce a self-sustaining nuclear fission chain reaction, where decay products stimulate other nuclei to decay. The more common Uranium-238 does not have this property.

[9]Technical note: I did these calculations using mathematical software that allows use of arbitrarily many decimal digits in floating point numbers rather than using a calculator, where the value of a would round to 1. I actually used 20-digit floating point numbers.

Therefore, solving as in part (a), we obtain

$$\log_{10}(a) = \frac{-\log_{10}(2)}{5.27} \doteq -0.05712.$$

Then

$$a \doteq 10^{-0.05712} \doteq .877.$$

To answer the question, then, we need to solve for t in the equation

$$(.05) \cdot Q(0) = Q(0)(.877)^t.$$

Taking logarithms, we find

$$t = \frac{\log_{10}(.05)}{\log_{10}(.877)} \doteq 22.8 \text{ years}.$$

We will reconsider what we did here next. △

The process we used in Example 5.8 parts (a) and (b) to solve for a is certainly direct. If we use some additional properties of logarithms, though, it is also possible to derive a formula giving a directly in terms of the half-life, so that this process can be bypassed (and the need for large numbers of decimal digits in computations in cases like part (a) can be avoided). Namely the definition of the half-life says if $t_{1/2}$ is that amount of time, then

$$(.5) \cdot Q(0) = Q(t_{1/2}) = Q(0)a^{t_{1/2}},$$

which shows

$$\log_{10}(a) = \frac{-\log_{10}(2)}{t_{1/2}}.$$

Instead of computing the value of the ratio on the right, we can also just consider the effect of raising 10 to both sides of the equation here. This yields

$$a = 10^{\log_{10}(a)} = 10^{\frac{-\log_{10}(2)}{t_{1/2}}} = \left(\frac{1}{2}\right)^{1/t_{1/2}}.$$

Hence the formula for an exponential decay model can be written in the following form *directly* if we know the half-life, $t_{1/2}$:

$$\boxed{Q(t) = Q(0) \cdot \left(\tfrac{1}{2}\right)^{t/t_{1/2}}.} \tag{5.9}$$

EXAMPLE 5.9. For example, a radioactive decay model for a quantity with half-life 12,000 years could also been written immediately using (5.9):

$$Q(t) = Q(0) \cdot \left(\frac{1}{2}\right)^{t/12000}$$

(with t in years). The corresponding value of a in the general form (5.3) is

$$a = \left(\frac{1}{2}\right)^{1/12000} \doteq .999653.$$

Note that this is slightly less than 1, as we expect for decay. △

5.4. Semi-log Plots and Fitting Exponential Models

In Chapter 4, we discussed the process of "fitting" a linear model to a data set using least squares regression. There is a corresponding process of fitting exponential models as well and we will introduce it this section. Via the "magic of logarithms" we will see, in fact, that all the techniques we learned before can be applied to this case as well!

If we are examining a data set and we notice a consistent curvature (or "concavity") in the point plot, then the residuals for a linear model might show the kind of consistent sign pattern we discussed in Chapter 4. If this happens, it is probably worthwhile to consider whether an exponential model is a better fit. A first step here might be to perform a standard transformation on the pointplot by plotting the dependent variable (function) values on a *logarithmic scale*. If the original data set consists of points (x_i, y_i), this would mean examining the new point plot showing the points $(x_i, \log_{10}(y_i))$ (any other logarithm function could also be used for this).

Example 5.10. A serious outbreak of the Ebola virus, primarily in the West African countries of Guinea, Liberia, and Sierra Leone, started in spring 2014 and continued into the fall of 2015. In the initial phases, many health care workers were infected by sick patients and died themselves.[10] Hospitals and clinics in the area were overwhelmed. In response, many countries outside the affected region sent medical assistance and a few of the volunteer medical workers who went were also infected, spreading a very small number of Ebola cases to the U.S., Spain, the U.K., and Italy in the fall of 2014. News coverage of these infections came close to causing panic at the time, even though the volunteer health worker victims were quickly quarantined by public health officials in their countries and there was no secondary spread outside the original African region. The following table shows the initial phase of the epidemic according to data from World Health Organization bulletins.[11] The word "cumulative" in the column headings means that the numerical values represent the total cases and deaths reported *up to and including the given date*. For instance there were 130 cases reported on or before 31 March, including the 49 cases reported on or before 22 March plus 81 new cases between 22 March and 31 March.

To investigate and model this data in the ways we have discussed, we need a way to convert the dates to numerical values. The simplest way to do that is to count the number of days after the initial date of March 22, yielding times in days as follows:

$$0 \quad 9 \quad 23 \quad 40 \quad 51 \quad 66 \quad 87 \quad 102 \quad 114 \quad 123 \quad 130.$$

We use these as the first coordinates (values of time as independent variable) to generate the plots in Figure 5.4. The points shown as black asterisks are the total cases; the red circles are the numbers of deaths. Examining this plot, it should be pretty clear that linear models would exhibit the sort of systematic sign patterns in the residuals that we warned about before. Because of the upward curvature, the

[10] Readers of Thucydides will recognize an eerie parallel here with the description of the plague that struck Athens in 430 B.C. in Book II of his history of the Peloponnesian war. Even with our better modern medical technology, outbreaks of disease cause many of the same problems that they did in the past.

[11] The data is summarized at `https://en.wikipedia.org/wiki/West_African_Ebola_virus_epidemic_timeline_of_reported_cases_and_deaths`, accessed June 30, 2017.

TABLE 1. Ebola Outbreak – March 22 through July 30, 2014.

Date	Cumulative Total Cases	Cumulative Deaths
22 March	49	29
31 March	130	82
14 April	176	110
1 May	239	160
12 May	260	182
27 May	309	202
17 June	528	337
2 July	779	481
14 July	982	613
23 July	1201	672
30 July	1437	825

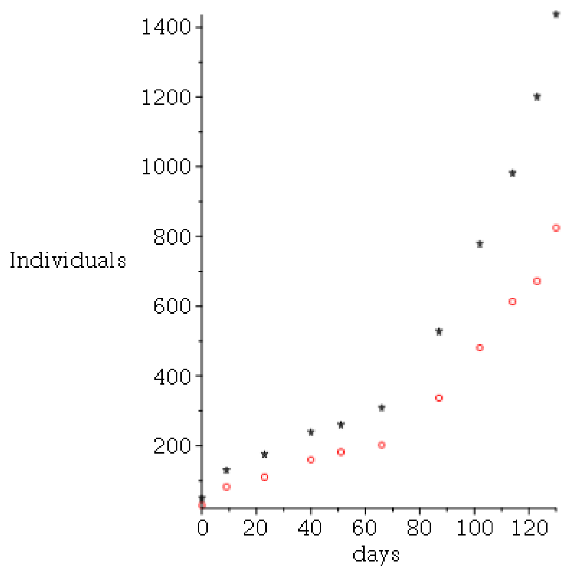


FIGURE 5.4. Ebola Outbreak, March 22 - July 31, 2014. Cases = $*$ (black), Deaths = $\circ$ (red).

first segment at the left of the point plot would lie above the regression line, the middle portion would lie below, and the final segment at the right would lie above

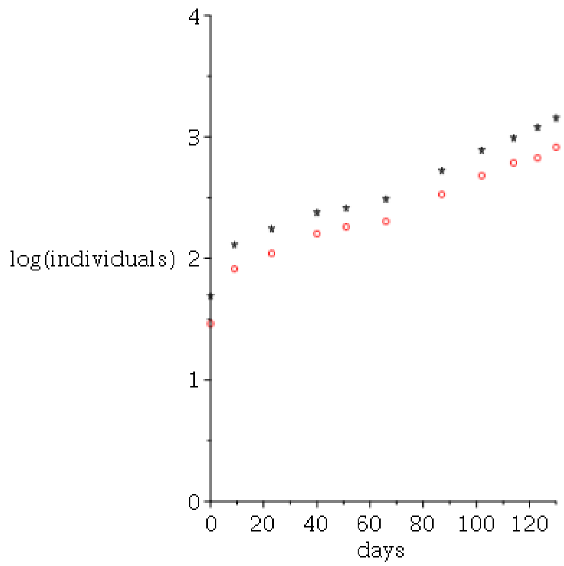


FIGURE 5.5. Ebola Outbreak, March 22 - July 31, 2014; vertical axis shown on logarithmic scale. Cases = $*$ (black), Deaths = $\circ$ (red).

again. Hence, it is possible that a different sort of model would match these data sets better.[12]

Let's see what happens if we take the logarithms of the numbers of infected individuals and deaths, shown in Figure 5.5. Note that (apart from an apparent anomaly in the data for $t = 0$, corresponding to March 22) these new plots look much more like what we expect if a *linear model* is going to be a good fit(!) The data points for March 22 also come from the very start of the epidemic; it is probably to be expected that the eventual pattern of spread characteristic of the early phase of the outbreak has not yet "set in." $\triangle$

Returning to the general discussion, plotting the points $(x_i, \log_{10}(y_i))$ is called generating a *semi-log plot* of the data set (the name is supposed to suggest that only the y_i and not the x_i are plotted on a logarithmic scale). If a linear model is a good fit for the transformed points shown in the semi-log plot then notice what happens. Suppose that

$$\log_{10}(y) = \hat{m}x + \hat{b} \tag{5.10}$$

is the fitted linear model and notice that we have written $\log_{10}(y)$ instead of y on the left hand side because of the data transformation we used. To obtain an equivalent model for the original quantity y, not $\log_{10}(y)$, we just raise 10 to the powers on both sides of (5.10):

$$y = 10^{\log_{10}(y)} = 10^{\hat{m}x+\hat{b}}.$$

[12]This is also indicated by various theoretical mathematical models of the spread of infectious diseases that we will discuss later in the course.

Using rules for exponents, this can be rewritten as

$$\boxed{y = (10^{\hat{b}}) \cdot (10^{\hat{m}})^x.} \tag{5.11}$$

This has the form given in (5.3) with

$$c = 10^{\hat{b}} \quad \text{and} \quad a = 10^{\hat{m}}.$$

Putting everything together we have the following process for fitting an exponential model to a dataset:

- Starting from the original (x_i, y_i) data points, apply the transformation $y_i \mapsto \log_{10}(y_i)$ to each point (and perhaps generate the semi-log plot of the points $(x_i, \log_{10}(y_i))$ to evaluate approximate linearity)
- Fit a linear model $\log_{10}(y) = \hat{m}x + \hat{b}$ to the transformed data set
- The corresponding exponential model for y as a function of x is obtained as in (5.11).

Example 5.11. We illustrate the process outlined above using the Ebola datasets studied in Example 5.10. The fitted linear model for the semi-log form of the cumulative total cases is

$$\log_{10}(y) = 1.915 + .0009543t,$$

(t in days) yielding

$$y = 10^{1.915} \cdot (10^{.009543})^t \doteq 82.2 \cdot (1.0222)^t.$$

Note that one way to interpret this model is that the number of infected individuals was increasing at a bit more than 2% *per day(!)* It is no wonder that the World Health Organization was immediately concerned about the severity of the outbreak.

Similarly, the fitted linear model for the semi-log form of the cumulative deaths is

$$\log_{10}(y) = 1.727 + .009299t,$$

(t in days again) yielding

$$y = 10^{1.727} \cdot (10^{.009299})^t \doteq 53.3 \cdot (1.0216)^t.$$

The cumulative deaths were increasing at a slightly slower rate, but still more than 2% per day. The models are plotted together with the datasets in Figure 5.6. The agreement between models and data is very close. △

To conclude our discussion, we mention two different ways to use the Excel spreadsheet environment to perform computations like these. You have the option of either

- doing the process as described and shown in this text, or
- using the *Exponential Trendline* options shown in the Chart Tools/ Trendline menu, but on the *untransformed*, (x_i, y_i) data set.

Those Excel options perform the computations exactly as we have described them but they automate the process so you don't have to do the logarithmic transformation yourself.[13] Also, Excel reports the result using the e^x exponential function, so some conversion is necessary to put the Excel output into the forms we have used.

[13] This means Excel is giving you a shortcut, but as always it's good to know what is going on "under the hood!"

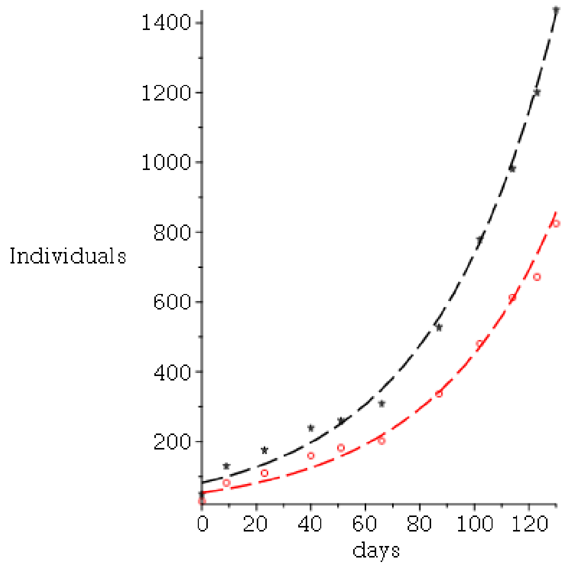


FIGURE 5.6. Ebola Outbreak, March 22 - July 31, 2014. Cases = $*$, Deaths = $\circ$, Models shown dashed.

Here's what you need to know. A function like $e^{.00133t}$ can be rearranged using rules for exponents:

$$e^{.00133t} = (e^{.00133})^t \doteq (1.00133)^t.$$

So this exponential model would be showing about 0.13% growth per unit time.

5.5. Chapter Project

The Chapter Project for this chapter will be to continue the study of the 2014-2015 Ebola outbreak begun above in Examples 5.10 and 5.11. The data we were using there comes from the Wikipedia page

`https://en.wikipedia.org/wiki/West_African_Ebola_virus_epidemic_timeline_of_reported_cases_and_deaths`

On that page, you will see a large table with the data we used at the bottom and rows above that giving the later course of the outbreak, all the way to November 2015, when the outbreak was declared to be concluded. (The rows are in reverse chronological order, for some reason.) The cumulative cases and deaths are also broken down by country, with separate totals for Guinea, Liberia, and Sierra Leone.

For the project, you want to start out by getting all of this data into Excel or another spreadsheet program. You will need to convert the dates to the numerical form we used above (days after the initial report of Ebola cases on March 22, 2014). It will be more convenient to have the rows in chronological order, too. There are a few values that are reported as ≥ 8881 or similar. You can just omit the $\geq$ for the purposes of this assignment.

Questions. Most of the work you will do will involve analyzing subsets of the data along the lines of what we did in the examples above, to try to develop answers to the following questions.

(A) Considering the total reported cases and total deaths for the whole course of the outbreak, for how long did the cases and deaths continue increasing at about 2% per day? When does it seem that the tide started to turn, in the sense that the epidemic was starting to "slow down" noticeably from the 2% per day rate of increase.[14] Try fitting models starting at March 22, 2014, but going farther than we did above. Some landmarks to watch for are November 2, 2014, 27 May 2015, 26 July 2015, and so forth. Explain how you are determining that "turn of the tide." Also, can you identify something definite that happened about that time that might have been a cause for the epidemic starting to come under control?

(B) Examining the data, it is pretty clear that the course of the outbreak was different in Guinea as compared to Liberia and Sierra Leone. Fit exponential models for initial phases of the epidemic in each of the three countries separately. Were the numbers of infections and deaths increasing at a rate of 2% per day at the start in all three countries?

(C) Not everyone who was infected with Ebola died as a result of the infection. Some people actually recovered. How did the percentage of infected people who recovered change over the course of the epidemic? Consider both the aggregate figures and the figures by country. What type of model fits that data the best? Note: One would hope that as an epidemic progresses, health care professionals would accumulate useful experience in how to treat the disease successfully so that their treatment outcomes would improve over time. Was that the case here?

Assignment. Submit a spreadsheet showing all of your calculations with the Ebola data set. Answer the questions above in a separate text document. As always, document your sources if you consult things other than the Wikipedia page given above.

Chapter Exercises

(1) We discussed an example of a family of circles with areas in a *geometric progression* in Figure 1.2. Formally, a geometric progression is any sequence of numbers of the form

$$c, c \cdot a, c \cdot a^2, c \cdot a^3, \ldots, c \cdot a^n$$

where c, a are fixed numbers. Show how such an geometric progression can be obtained as the values $f(0), f(1), \ldots, f(n)$ for some exponential function. (What is the formula of that exponential function?)

(2) Solve the following exponential equations using logarithms proceeding as in Example 5.2.

[14] It had to do that at some time; if it didn't, then we would probably all be dead by now!

(a) $28.3 = 4.5(3.4)^{2x}$.

(b) $4^x = 3.5 \cdot 2^x$.

(c) $7.9 = 2.8(7.4)^x + 5.6$.

(3) Find the exponential functions $f(t) = c \cdot a^t$ satisfying:

(a) $f(0) = 3.2$ and $f(1) = 2.9$.

(b) $f(0) = 7.8$ and $f(5) = 4.3 \cdot f(1)$.

(c) $f(2.1) = 20.3$ and $f(5.6) = 25.4$.

(4) Find the exponential functions $f(t) = c \cdot a^t$ satisfying:

(a) $f(0) = 8.54$ and f decreasing at a 15% rate for each unit change in t. (Hint: See (5.8).)

(b) $f(1.2) = 3.7$ and f growing at a 3% rate for each unit change in t.

(5) Redo the calculations of Example 5.8 but using (5.9) instead of computing a from the half-life.

(6) Let $Q(t)$ be an exponential function of time t. The *doubling time* of Q is the value $t = t_2$ for which $Q(t_2) = 2Q(0)$.

(a) What is the doubling time for $Q(t) = 5 \cdot (1.25)^t$?

(b) What is the doubling time for $Q(t) = 18.3 \cdot (3.4)^t$?

(7) Show in general that the doubling time for any $Q(t) = Q(0)a^t$ is given by

$$t_2 = \frac{\log_{10}(2)}{\log_{10}(a)}.$$

What does this mean if $0 < a < 1$?

(8) Show that an exponential model with a doubling time t_2 can be written down immediately in the form

$$Q(t) = Q(0) \cdot 2^{t/t_2}.$$

What is the corresponding value for a as in (5.3)? (Hint: Look at the derivation of (5.9) in the text.)

(9) The radioactive isotope Strontium-90 has a half-life of 28.8 years and is produced by nuclear fission (in particular, it was produced by early nuclear weapons testing). It has associated human health concerns because it is chemically equivalent to stable Strontium-88 and hence similar to Calcium. It tends to be concentrated in the bodies of cows ingesting contaminated plant material and it is excreted in their milk.

(a) Determine an exponential model $Q(t) = Q(0)a^t$ for the amount of Strontium-90 present in a sample as a function of time.

(b) How long will it take for a sample of Strontium-90 to decay to .01% of the original amount present?

(10) The radioactive isotope Carbon-14 has a half-life of about 5730 years. It is naturally present in the atmosphere at the level of 1 part per trillion, almost entirely as a product of an interaction between cosmic rays and Nitrogen atoms in the

upper atmosphere. Carbon-14 is chemically similar to stable Carbon-12, and hence forms compounds that are incorporated into plants and animal tissues. As long as the plant or animal is alive, the Carbon-14 in its tissues is replenished by its metabolic processes from the reservoir of atmospheric Carbon-14. However, as soon as the plant or animal dies, the amount of Carbon-14 remaining in the organic materials making up the remains (and hence in objects that might have been made by humans from those remains) begins to decay. For this reason, measurements of the ratio between Carbon-14 and Carbon-12 can be used to estimate the age of those organic materials (i.e. the time elapsed since the death of the plant or animal they came from).[15]

(a) The rate of production of Carbon-14 has been estimated to be about 1.7 atoms of Carbon-14 per square centimeter, per second over the whole surface area of the Earth.[16] How many atoms of Carbon-14 is that per year?

(b) Determine an exponential model $Q(t) = Q(0)a^t$ for the amount of Carbon-14 present in a sample as a function of time.

(c) A piece of cloth found in a newly discovered tomb has a Carbon-14-Carbon-12 ratio 34% as large as the Carbon-14-Carbon-12 ratio in a comparable newly-made piece of cloth. Use this information to estimate how old the tomb is. What assumptions are you making?

(11) We discussed various examples of human population data in Chapter 4 and fitted linear models. In this exercise, you will work with the historical record of U.S. population data from the Federal Census taken every 10 years given in Table 2.

(a) Fit linear and exponential models to this data set. Is one of them a noticeably better fit than the other? Interpret what your model says about the rate of growth of the U.S. population over time.

(b) The discussion of population models in Example 5.1 involves several assumptions that are not good matches for the situation with the U.S. population. For example, has the U.S. population changed only by births and deaths? Have the birth and death rates been *constant*? Discuss and answer by referring to the data in Table 2.

(12) The data in Table 3 (two pages ahead) gives the length (in mm) and the weight (in g) of 42 fish caught in the Spokane River in Washington State.[17] (Each row of the table gives the data for two fish.)

(a) Fit linear and exponential models to this data set. Is either one clearly superior?

---

[15] This *radiocarbon dating* technique was developed by a group led by a physical chemist named Willard F. Libby (1908 - 1980). Libby devoted part of his career to work with radioactive isotopes–including a role in a branch of the World War II "Manhattan Project" that produced the first usable fission nuclear weapons–and part to environmental issues including development of air pollution standards in California. He received the 1960 Nobel Prize in Chemistry for the radiocarbon dating technique.

[16] See for instance, G. Kovaltsov, A. Mishev, I. Usoskin, "A new model of cosmogenic production of radiocarbon ^{14}C in the atmosphere," arXiv:1206.6974.

[17] Source: Langkamp and Hull, Quantitative Environmental Learning Project, `http://seattlecentral.edu/qelp/sets/023/023.html`, downloaded on June 30, 2017.

Table 2. U.S. Population According to Federal Census Records

Census	Population	Census	Population
1790	3,929,214	1800	5,308,483
1810	7,239,881	1820	9,638,453
1830	12,866,020	1840	17,069,453
1850	23,191,876	1860	31,443,321
1870	38,558,371	1880	50,189,209
1890	62,979,766	1900	76,212,168
1910	92,228,496	1920	106,021,537
1930	123,202,624	1940	132,164,569
1950	151,325,798	1960	179,323,175
1970	203,211,926	1980	226,545,805
1990	248,709,873	2000	281,421,906
2010	308,745,538		

(b) (Looking ahead to Chapter 6.) Does either a linear or exponential relationship between length and mass seem likely on geometric grounds? For instance, let us consider a (very) simplified model where the fish are all *scaled copies of one another*, with tissues of constant mass density (mass per unit volume). That is, suppose the individual fish have exactly the same shape and uniform density. But for each individual fish, each linear dimension in the "standard model" of this species is multiplied by a constant[18] to get the corresponding dimensions of the individual. What would be true about the mass as a function of length in that case?

[18] "Scaled copy" would mean these constants are the same for all dimensions in each individual fish, but they could vary from individual to individual.

TABLE 3. Trout weight and length data.

Length (mm)	Weight(g)	Length(mm)	Weight(g)
457	855	405	715
455	975	460	895
335	472	365	540
390	660	368	581
385	609	360	557
346	433	438	840
392	623	324	387
360	479	413	754
276	235	387	538
345	438	395	584
326	353	270	209
359	476	347	432
259	202	247	184
280	248	265	223
309	392	338	460
334	406	332	383
324	353	337	363
343	390	318	340
305	303	335	410
317	335	351	506
368	605	502	1300

CHAPTER 6

Power Functions as Models

6.1. Introduction

In the two previous chapters we have studied using linear and exponential functions as models. In this chapter we turn to another type of model based on *power functions*, or functions of the form

$$\boxed{y = \beta \cdot x^{\alpha},} \tag{6.1}$$

where x is the independent variable, y is the dependent variable, and α and β (the Greek letters "alpha" and "beta") are constants. These are used quite commonly in biology, ecology, and other scientific areas, both as theoretical models and as tools to analyze experimental data. We will introduce the idea with the following example.

Example 6.1. Consider the question of how many individuals of a given animal species can coexist in some fixed environment, say a fixed area, and how that might depend on the typical body sizes of the animals. Clearly, the larger the animals' bodies are, the smaller the number the area can support should be. But can we say more?

The individual animals will have some average resource use rate per unit time, R, while the environment will produce those resources at some rate per unit area per unit time, P. For instance, if the species is herbivorous, the R would represent how much plant biomass an individual consumes per unit time, while P would represent the rate of production of that plant biomass per unit time by the environment. Similarly, if the species is carnivorous and predatory, R would represent how much animal biomass an individual consumes per unit time, while P would represent the rate of production of the prey species the predators consume for food per unit area per unit time. In either case, the *maximum number of individuals* that could be supported, which we will denote by N, would be proportional to the ratio between the resource productivity of the environment and the individual resource use:

$$N = C \cdot \frac{P}{R}, \tag{6.2}$$

where C is some constant.

It is clear that R, the average resource use rate for the individual animals, is correlated with their average body mass, since larger bodies require more food intake to maintain their metabolisms. By examining this relationship, biologists have noted the empirical observation that R tends to be proportional to the body mass M raised to the $3/4$ power:

$$R = C_0 \cdot M^{3/4} \tag{6.3}$$

for some constant C_0. A closely related statement is known as *Kleiber's Law* – the metabolic rate in animal bodies should follow a power law with

$$\text{metabolic rate} = C \cdot M^{3/4}.$$

We will examine a real-world data set that is related to this pattern later in the chapter. For now, note that we have an example of a power function as in (6.1) with $\alpha = 3/4$. This says, in particular, that the resource use rate does not increase *linearly* with M, but somewhat less than linearly.

If we substitute for R from (6.3) into (6.2), then we see

$$N = \frac{C}{C_0} \cdot M^{-3/4},$$

which expresses the maximum sustainable population N as another power function of the body mass M. Note that this one has exponent $\alpha = -3/4 < 0$. This indicates that N should decrease as M increases, as is only reasonable from our intuition about the real world situation. But note that it gives a much more precise mathematical model describing a theory about how the relation between body size and sustainable population should work! With experimental data we could then either find evidence to support the theory or perhaps indicate that it needs revision or modification. △

In this chapter, we will consider power functions in general first, then look at their use in modeling. We will introduce some techniques for fitting power law models to data and consider a number of examples.

6.2. Power Functions

Power functions are those given by the general form in (6.1). The constant α is sometimes called the *scaling exponent* because scaling x by a constant factor c – that is, replacing x by cx multiplies the output value y by the factor c^α. In the applications we usually restrict to $x \geq 0$, and these inputs are *always* in the domain of a power function. The $x < 0$ will not be in the domain of power functions such as $x^{1/2}$. So restricting to $x \geq 0$ simplifies the discussion somewhat.

Example 6.2. We have encountered several simple power functions in examples in previous chapters. For instance, the area of a square of side x is $A = x^2$, a power function with exponent $\alpha = 2$. The volume of a cube of side x is $V = x^3$, a power function with exponent $\alpha = 3$. These geometric formulas are the reason for calling x^2 "x squared" and x^3 "x cubed." The area of a circle as a function of the radius is given by $A = \pi r^2$, which is another power function with exponent $\alpha = 2$. △

The graphs of power functions have several forms. In Figure 6.1, we plot several power functions with the scaling exponent $\alpha > 0$. The solid red curve is $y = x^{1/2}$; the dashed blue curve is $y = x^{3/4}$, the solid black curve is the linear function $y = x$ (linear functions with intercept $= 0$ are special cases of power functions); the dashed red curve is $y = x^2$; the solid blue curve is $y = x^3$; the dashed black curve is $y = x^{4.3}$. Note that:

- The graphs of the functions with $\alpha > 1$ are *concave up* for all $x \geq 0$ and cross from below the line $y = x$ to above that line at $x = 1$.

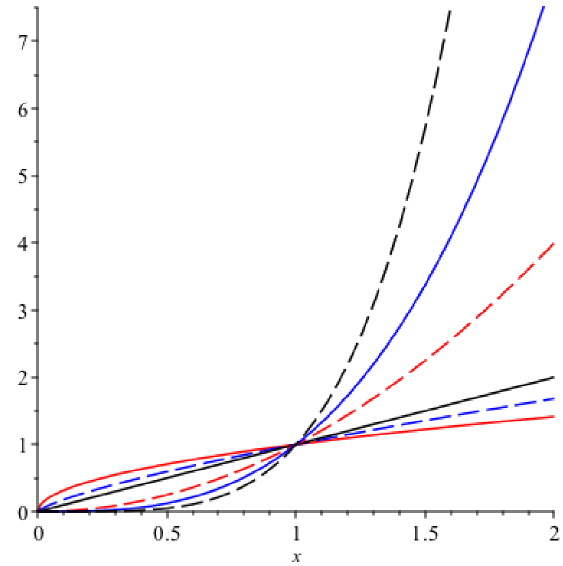


FIGURE 6.1. Power functions $y = x^{\alpha}$ with $\alpha > 0$

- The graphs of the functions with $0 < \alpha < 1$ are *concave down* for all $x \geq 0$ and cross from above the line $y = x$ to below that line at $x = 1$.
- The graph of the function with $\alpha = 1$ is a *straight line* and something like a *boundary case* between the other two types of power law graphs with $\alpha > 0$.

All of the plots pass through the point $(1, 1)$ because we plotted functions with the constant $\beta = 1$ in the general form from (6.1) in all cases. Note that $f(x) = \beta \cdot x^{\alpha}$ would satisfy $f(1) = \beta$ instead. Changing β to some value other than $\beta = 1$ multiplies all function values by that constant and stretches or shrinks the whole graph vertically depending on whether $\beta > 1$ or $0 < \beta < 1$.

The power functions with a negative scaling exponent are somewhat less variable. In Figure 6.2, the solid curve is $y = x^{-2}$, the dashed curve is $y = x^{-1}$ and the dotted one is $y = x^{-1/2}$. No power function with $\alpha < 0$ is defined at $x = 0$, and each of the graphs has a *vertical asymptote* there. As x approaches 0 through positive values, the value of each of these functions grows without bound. The larger $|\alpha|$ is, the greater the values of the function for $0 < x < 1$ are. But the pattern reverses for $x > 1$. There, the larger $|\alpha|$ is, the smaller the values of the function are for $x > 1$.

Power Functions Versus Exponential and Logarithm Functions. There are some superficial similarities between certain power functions, exponentials, and logarithms that should be noted. But these functions should not be confused with one another. First, be careful that you understand the difference between powers and exponentials. Even though the values of both functions can often be computed with the same calculator button (the ˆ button on many scientific calculators), these are quite different. In the a^x exponential, it is the base that is constant and the

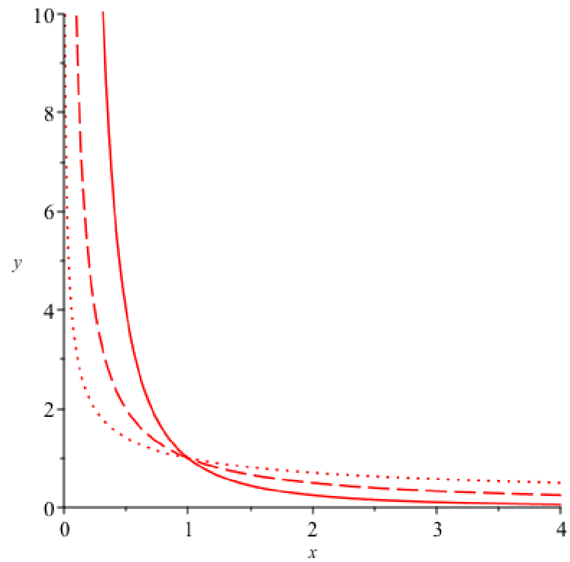


FIGURE 6.2. Power functions $y = x^\alpha$ with $\alpha < 0$

variable is in the exponent. On the other hand, in the power function x^α, it is the exponent that is constant and the base is the variable.

Power functions x^α with $\alpha > 1$ and exponentials a^x with $a > 1$ also have somewhat similar-looking graphs – they are both increasing and "concave up." If only a subset of the domain is considered, it can appear as though a power function is growing faster than an exponential. For instance, in Figure 6.3a, the solid curve is $y = x^3$, while the dashed graph is $y = 2^x$ and the interval of x-values shown is $0 \le x \le 3$. Here it looks as though $y = x^3$ is growing faster. On the other hand, if we extend the plotting interval to $0 \le x \le 11$, as in Figure 6.3b, then we can see that the exponential $y = 2^x$ has overtaken $y = x^3$ around $x = 10$ and is growing more quickly after that point.[1] In fact, it can be shown, using techniques from calculus that:

- *Every exponential function* a^x with $a > 1$ eventually grows faster than every power function x^α, in the sense that $\frac{x^\alpha}{a^x}$ eventually tends to zero as x increases without bound.[2]

The graphs of power functions with $0 < \alpha < 1$ have a superficial similarity to graphs of logarithm functions. Both are increasing and "concave down" as in Figure 6.4. There, $y = x^{1/2} = \sqrt{x}$ is shown solid and $y = \log_2(x)$ is shown dashed.

---

[1]Technical note: $2^{10} = 1024$, while $10^3 = 1000$, so by $x = 10$, the exponential has already exceeded the power and the curves cross for x slightly less than 10. There is no simple algebraic way to solve the equation $2^x = x^3$, though, because the unknown x appears both in the exponent and in the base of the power. Root-finding tools based on calculus are necessary to approximate the solutions. There are two solutions visiible in Figures 6.3a and 6.3b. The values are approximately $x = 1.373467120$ and $x = 9.939535141$.

[2]Technical note: In calculus terms, we are saying that when $a > 1$, $\lim_{x\to\infty} \frac{x^\alpha}{a^x} = 0$ no matter how large α is. This can be shown using L'Hopital's Rule.

(A) Plotted on $0 \leq x \leq 3$. (B) Plotted on $0 \leq x \leq 11$.

FIGURE 6.3. The exponential 2^x (dashed) grows faster than the power x^3 (solid) for large enough x.

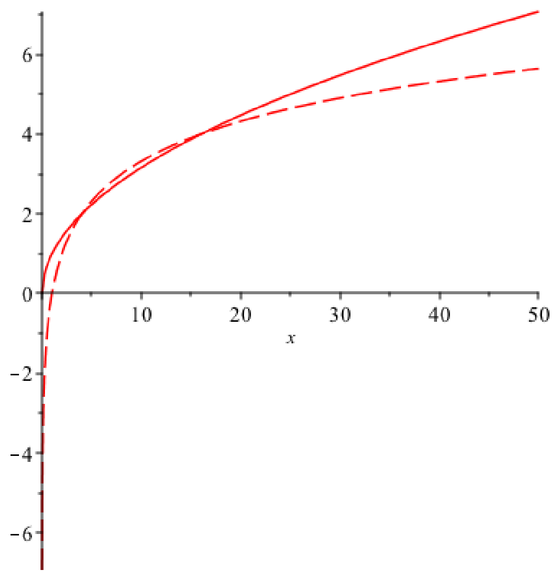


FIGURE 6.4. Power function $y = x^{1/2}$ (solid) and logarithm $y = \log_2(x)$ (dashed).

However, the power function graph does not cross the x-axis at $x = 1$ and it does not have the portion below the x-axis for $0 < x \leq 1$. On that range, the logarithm is approaching a vertical asymptote at $x = 0$ (with y going to $-\infty$ as x goes to zero).

Working With Power Functions. Since there are two parameters α and β in the general form (6.1), given two distinct points in the first quadrant of the plane (that is, with $x, y > 0$), there is exactly one power function whose graph passes through those two points.

EXAMPLE 6.3. Say the two points are $(2.3, 5.6)$ and $(3.4, 2.1)$. We want to find a power function $f(x) = \beta x^\alpha$ such that

$$\begin{aligned} 5.6 &= \beta \cdot (2.3)^\alpha \\ 2.1 &= \beta \cdot (3.4)^\alpha. \end{aligned} \tag{6.4}$$

If we divide the first equation by the second, then canceling the β factors gives

$$2.667 \doteq \frac{5.6}{2.1} = \frac{\beta(2.3)^\alpha}{\beta(3.4)^\alpha} = \left(\frac{2.3}{3.4}\right)^\alpha \doteq (.6765)^\alpha.$$

Taking logarithms (to solve for the α in the exponent!) we get

$$\log_{10}(2.667) = \alpha \cdot \log_{10}(.6765) \Rightarrow \alpha \doteq -2.51.$$

Then we can solve for β using either of the equations in (6.4). Using the first gives

$$5.6 = \beta \cdot (2.3)^{-2.51} \Rightarrow \beta = 5.6 \cdot (2.3)^{2.51} \doteq 45.3.$$

The power function we want is (approximately) $f(x) = 45.3 \cdot x^{-2.51}$. It is easy to check then that $45.3 \cdot (3.4)^{-2.51}$ is approximately equal to 2.1. Note also that since $5.6 > 2.1$, we must have that the power function whose graph passes through these two points is decreasing. Hence we expect a *negative* α value by looking at Figures 6.1 and 6.2. The value $\alpha \doteq -2.51$ matches this. $\triangle$

6.3. Log-Log plots and Fitting Power Law Models

We will say one quantity y follows a *power law* if y is given by a power function of some other quantity x, that is

$$y = \beta \cdot x^\alpha$$

as in (6.1). Any equation of this form can also be thought of a mathematical model describing how the quantities y and x are related.

As with the exponential models studied in Chapter 5, the most basic standard process of fitting a power law model to a data set proceeds by applying the least squares regression technique to a transformed data set. We can see the idea if we start from (6.1) and take logarithms on both sides:

$$y = \beta \cdot x^\alpha \Leftrightarrow \log_{10}(y) = \log_{10}(\beta \cdot x^\alpha) = \alpha \log_{10}(x) + \log_{10}(\beta),$$

using parts (1) and (4) of Proposition 1.8. Note that we have applied the logarithm function to *both x and y* here. We will call this the *log-log transform* on the original data set, and the resulting point plot is called a log-log plot. Note that this means *both* the independent and dependent variables are plotted using a logarithmic scale as discussed in Chapter 1. From this we see:

- *If the data set (x_i, y_i) follows a power law $y = \beta \cdot x^\alpha$ exactly, then the log-log transformed data points $(\log_{10}(x_i), \log_{10}(y_i))$ lie on the straight line with slope α and intercept $\log_{10}(\beta)$.*
- *Even if not, a power law model can be fit to the data set by finding the best linear fit for the log-log transformed data set using least squares regression. If the best fit line has slope $\hat{m}$ and intercept $\hat{b}$, then $\alpha = \hat{m}$ and $\beta = 10^{\hat{b}}$ are the parameters of the best fit power law.*

In "the old days," both this process and the process described in Chapter 5 for fitting exponential models were often carried out using specially-designed graph paper. Semi-log graph paper has the vertical axis ruled on a logarithmic scale; log-log graph paper has both axes ruled this way. One very basic and approximate way to carry out these computations is to plot the data set on the appropriate sort of graph paper, estimate the "best-fit" line *visually*, and use that to derive the appropriate exponential or power law model. It is still possible to find templates for printing out both kinds of graph paper online today, but we will not use them because estimating the best-fit line that way tends to be a somewhat ambiguous and subjective process. We will use Excel or other software to carry out the calculations more precisely.

Just as for the exponential models from Chapter 5, there are two different ways to use the Excel spreadsheet environment to perform computations like these. You have the option of either

- doing the process as described and shown in this text, or
- using one of the *Trendline* options shown in the Chart Tools/Trendline menu, namely the *Power* option, but on the *untransformed*, (x_i, y_i) data set.

Those Excel options perform the computations exactly as we have described them above (by the log-log transform) but they automate the process so you don't have to do the log-log transformation yourself.

We next illustrate the process for fitting a power law model on an example.

Example 6.4. The data in Table 1 shows a range of typical resting heart rates (in beats per minute – bpm) for adults of a number of mammalian species and an adult body mass value for the same species.[3] The body mass values here are

Table 1. Mammalian Resting Heart Rate and Body Mass Ranges

Species	HeartRate (bpm)	BodyMass (g)
cat	$120 - 140$	$2,900$
cattle	$48 - 84$	$347,000$
dog	$70 - 120$	$40,000$
Asian elephant	$25 - 35$	$2,720,000$
goat	$70 - 80$	$33,500$
guinea pig	$200 - 300$	720
horse	$28 - 40$	$250,000$
rabbit	$180 - 350$	$1,580$
human	$60 - 84$	$70,000$

estimated averages. We will think of the body mass as the independent variable here and the heart rate as the dependent variable. For simplicity, we use the midpoint of each range of heart rates.[4] Applying the log-log transform, we find the straight line best fitting the transformed data set and we find

$$\log_{10}(\text{heart rate}) = -0.265 \cdot \log_{10}(\text{body mass}) + 3.146.$$

[3]The non-human heart rate ranges are from S. Fielder, "Resting Heart Rates," `http://www.merckvetmanual.com/appendixes/reference-guides/resting-heart-rates`, average body masses from `eol.org`, both accessed July 2, 2017.

[4]For some of the species, for instance for *Canis lupus familiaris* (the domestic dog), there is a tremendous range of body sizes, weights, and heart rates. Hence we try to average over the whole range of the species in making this comparison.

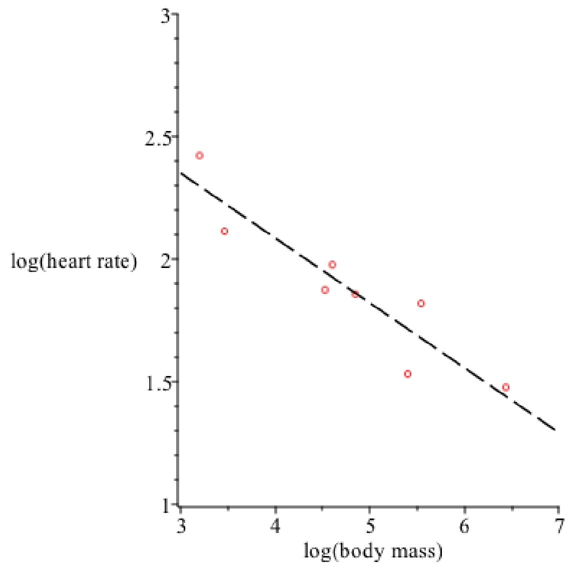


FIGURE 6.5. The log-log transformed point plot of the body mass/heart rate data, and best fit line.

Raising 10 to both sides of this equation we get the fitted power law model,

$$\text{heart rate} = 1400 \cdot (\text{body mass})^{-.265}.$$

The log-log transformed point plot and the regression line are shown in Figure 6.5. The question we are considering here is related to the relation between body mass and metabolic rate (Kleiber's Law) discussed in the Introduction. This relationship between body mass and resting heart rate is thought to follow a power law with scaling exponent $\alpha = -1/4$.[5] The results from the analysis of our small data set agrees with that pretty well! $\triangle$

For a second example, we will consider the data set of measured lengths and masses of trout from Exercise 12 in Chapter 5.

EXAMPLE 6.5. In Exercise 12 from Chapter 5, you fit linear and exponential models to the data set from Table 3. Figure 6.6 shows the point plot of the "raw" data set (i.e. before applying any logarithmic transformations). We view the length as the independent variable and the mass as the dependent variable. Visually, we can see that a linear model would be reasonably good here, although there is a certain amount of upward curvature (concavity upward) in the point plot as the length variable increases and this would tend to produce a sign pattern in the residuals for the linear regression. Hence an exponential model is also a possibility here, and computing R^2 values indicates that the exponential model is perhaps slightly better ($R^2 \doteq .94$ for the linear model versus $R^2 \doteq .96$ for the exponential model).

[5] V.M. Savage, et. al. "The predominance of quarter-power scaling in biology," Funct. Ecol. 2 April, 2004.

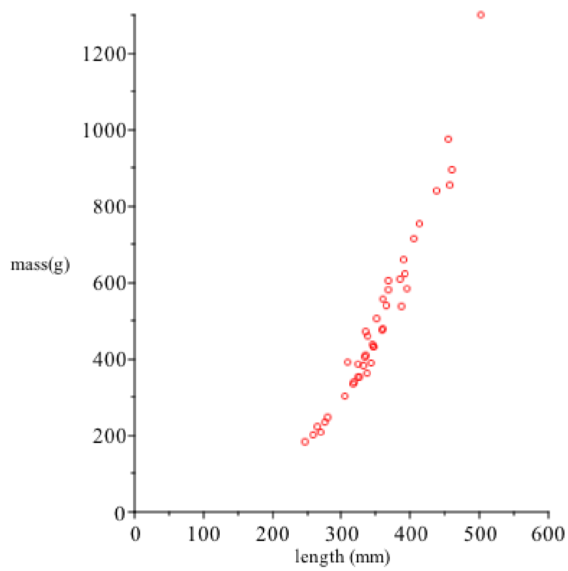


FIGURE 6.6. The data from Table 3 of Chapter 5.

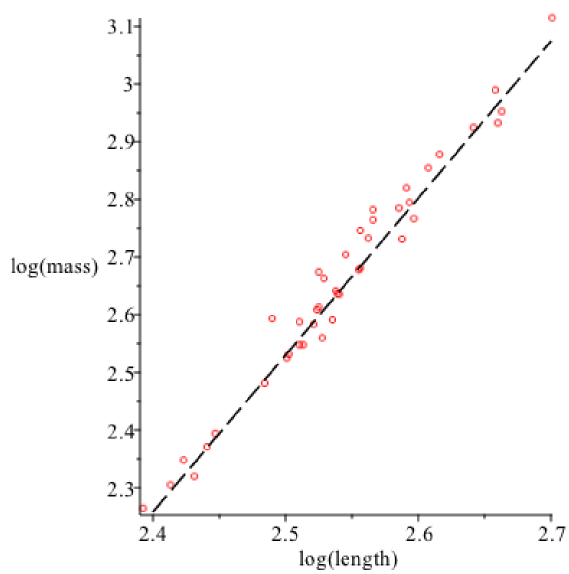


FIGURE 6.7. Log-Log transform of trout data and regression line.

Let us also fit a power law model. The log-log transformed data set and the regression line are shown in Figure 6.7. The equation of the regression line is

$$\log_{10}(\text{mass}) = 2.724 \cdot \log_{10}(\text{length}) - 4.268.$$

So our power law model is

$$\text{mass} \doteq (5.395 \times 10^{-4}) \cdot (\text{length})^{2.724}.$$

The R^2 statistic here is $R^2 \doteq .97$, which indicates that the power law is perhaps a slightly better fit than either the linear or the exponential model (these are not large differences though and we should not read too much into them). Note that the scaling exponent 2.724 is also relatively close to the value of 3 that you should have derived from the simplifying assumption that the fish are all scaled copies of each other in part (b) of Exercise 12 from Chapter 5. A possible explanation for the fact this is not closer to that theoretical value of 3 is that the fish are not just scaled copies of one another (i.e. for instance, knowing that one fish is 1.2 times the length of another does not mean that every structure in the two bodies has dimensions in the same proportion). Moreover, different tissues in the fish bodies have different mass densities and variations in scaling between different structures may produce deviations from a power law with scaling exponent 3. $\triangle$

This last example shows an important consideration to keep in mind. Namely, the techniques we have introduced for fitting models of various sorts (linear, exponential, power law) can be used in different ways in different circumstances. In some situations, there is a "preferred" type of model to use based on a theoretical understanding of what is happening in the real-world situation being studied. For example, over a short time frame, even though a linear model might fit measured amounts of a radioactive isotope undergoing decay quite closely, we also know on theoretical grounds that the exponential models are a better description of the underlying physical process. So exponential models would be preferred and the goal might be to estimate the half-life of the isotope from the fitted model. Similarly, in the situation of Example 6.5, consideration of the geometry of the fish bodies seems to indicate that a power law would be a better model for how weight depends on length than either a linear model or an exponential model on theoretical grounds and then the question becomes: What scaling exponent matches the data for a given species? But in other cases, we may not have a theoretical model in hand and several or all of the sorts of models we have discussed may be applicable. As we have discussed, there are indications to look for that show a linear model might not be matching a data set well, and then exponential or power law models might yield better results.

On the other hand, you should be cautious about attempting to *infer a theoretical exponential or power law relationship between quantities on the basis of the results of modeling* using the techniques we have discussed. This is a much more subtle problem and it calls for more sophisticated statistical tools than we will be able to consider in this introductory treatment.[6] We will see some examples that indicate the subtleties involved in the next section.

6.4. Power Law Distributions

To discuss our next application of power law models, we need to begin by introducing another data analysis concept. It is often of interest to understand how the items in a data set are distributed. One way to analyze this is to construct

[6] There are times when consulting an expert statistician is necessary!

a *frequency histogram*, which shows how many of the items in the given data set lie in each one of a given set of ranges of values, or *"bins."* Histograms are often plotted as bar charts as in Figure 4.2 from Chapter 4.

EXAMPLE 6.6. Consider the following hypothetical (i.e. "made-up") data set with $N = 17$ values:

$$\begin{array}{ccccccccc} 12.4 & 5.6 & 21.3 & 4.3 & 23 & 24.4 & 2.3 & 17.6 & 12.0 \\ 13.9 & 21 & 27.0 & 12.2 & 14.1 & 3.4 & 11.8 & 12.9. & \end{array} \tag{6.5}$$

We see that all of the data lies in the numerical range 0.0 to 28.0. If we divide this range into 7 equal subintervals, or *"bins"* of width 4, then the frequency histogram is based on the number of data values that fall into each of the bins, as in Table 2. The value 12.0 comes at the boundary between two of the bins. As indicated, we counted it in the bin for $12 \le x < 16$, not the bin for $8 \le x < 12$. Note that

TABLE 2. Frequencies in data from (6.5)

Bin	Frequency
$0 \le x < 4$	2
$4 \le x < 8$	2
$8 \le x < 12$	1
$12 \le x < 16$	6
$16 \le x < 20$	1
$20 \le x < 24$	3
$24 \le x < 28$	2

the entries in the Frequency column sum to $N = 17$ since we have accounted for all of the data values. The corresponding histogram is then a bar chart in which the heights of the bars represent the Frequency values. These might be plotted, for centered at the midpoints of the bins, as in Figure 4.2 from Chapter 4. See Figure 6.8 for the frequency histogram bar chart from this data. $\triangle$

There are many choices involved in the constructions of histograms:

- how many bins are used,
- where the bin boundaries are placed and what the widths of the bins are,
- how the values at the boundaries of bins are handled.

In our example we arbitrarily chose 7 bins of width 4 and handled boundaries by saying that any values at a boundary between two bins would count in the higher bin. For instance, the bin definition $0 \le x < 4$ in Table 2 means that if we had a value of 4.0, it would count in the next bin for $4 \le x < 8$, not in the bin for $0 \le x < 4$. It would also always be possible to (re)define bins so that no data values occur at bin boundaries(!)

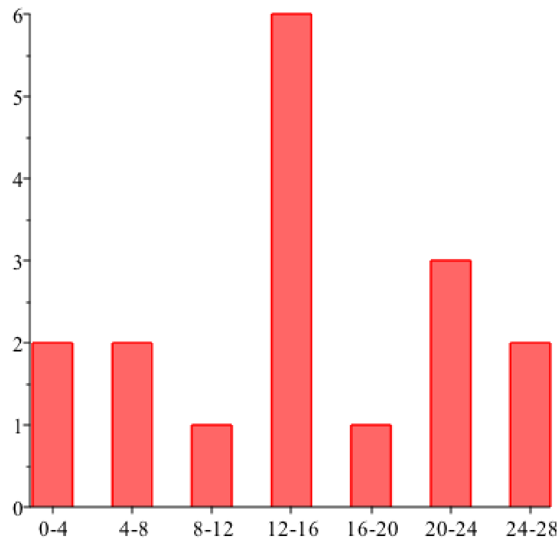


FIGURE 6.8. The frequency histogram for the data from Table 2.

All of this means that it is not possible to specifiy any one histogram that is the sole correct representation of the distribution of values in a data set.[7] In general, however, rules of thumb that are usually followed include:

- Unless there are clear *outliers* far outside the range of values in the rest of the data, the bins should cover the whole range of values represented in the data set.
- Somewhere between 5 and 10 bins are normally used; using fewer might mean too many data values may get "lumped" into one bin for important underlying structure to be revealed, while using more might mean that not enough aggregation into groups happens – missing the forest for the trees, so to speak.
- It is not necessary for all the bins to have equal width, but in that case, it is "good form" to plot the bar chart using rectangles the whole width of the bin to show how the bins are related.

A variant of the frequency histogram is the *relative frequency* histogram where the heights of the bars are scaled by dividing by the number of data values, N, and (possibly) multiplying by 100%. If the heights are converted to percents, then the bars in a histogram represent the portion of the data set contained in the different bins. For example, in the relative frequency histogram corresponding to Figure 6.8, the bar for the bin $12 \le x < 16$ would have height $6/17 \times 100\% \doteq 35.3\%$, so about 35% of the data values are in that range.

For large data sets, it can be quite tedious to count the numbers of data values in each bin to construct a frequency histogram. Hence software like Excel

[7]There are also many "shady" tricks that can be used to conceal patterns in data or create the appearance of patterns that are not present. As Mark Twain was fond of saying, "there are three kinds of lies: lies, damned lies, and statistics."

has facilities to automate the process. The controls for doing this are under the Tools/Data Analysis pull-down menu. (If you do not see Data Analysis under the list of Tools options, you will need to install the *Analysis ToolPak* add-on package. This is always distributed with Excel, but not all users need its features, so it is not automatically installed when you launch Excel for the first time. Follow the instructions under the listing for the *Analysis ToolPak* to install it.) Once you have the Tools/Data Analysis going, the steps to construct a histogram are the following:

- Enter your data in one range of cells, and a list of *upper bin boundaries* in a second range of cells.
- Select Histogram from the list of Tools/Data Analysis functions.
- Enter the range of cells containing your data.
- Enter the second range of cells where you have entered the bin boundaries.
- Enter the options you want (indicate whether you want the histogram plotted, and whether you want the output on a separate page of the spreadsheet).
- Press the OK button.

We now discuss how histograms and the power law models we have introduced can be used to understand the distribution of many real-world data sets. Many kinds of real-world data have frequency histograms showing that *small values of some quantity are much more common than large values.* For instance, small towns are much more numerous than larger cities. Small wildfires that burn an acre or two are much more common than the massive fires that burn many square miles of forest. Small oil spills are much more common than the "mega-disasters" like the Deepwater Horizon oil well blowout in the Gulf of Mexico in 2010.

Power law models can be used to describe the distribution of quantities like the population of cities, or the area burned in wildfires, or the amount of oil spilled in accidents in some situations. A standard method for fitting a power law model to a distribution is as follows

- Construct a histogram for the data
- For each bin, start by computing what is called the *reverse cumulative frequency* – the number of data points with values *greater than or equal to the lower boundary of the bin.*
- Fit a power law model to the data set (x_i, y_i) where x_i = midpoint of ith bin, and y_i is the reverse cumulative frequency for the ith bin.
- If the power law model (which will almost always have $\alpha < 0$) fits the data well (R^2 close to 1, or R close to -1), then we say the data has an (approximate) *power law distribution.*

We apply this idea to study the distribution of several data sets.

EXAMPLE 6.7. According to the 2010 Federal Census, the 127 largest cities and towns in the state of Massachusetts had populations distributed as in Table 3. The one extremely large value was Boston, with a population of $617,594$; the cities of Worcester and Springfield were the unique cities in the $160,000 - 200,000$ and $120,000 - 160,000$ bins, respectively. Figure 6.9 shows the log-log plot of the reverse cumulative frequencies versus the bin midpoints, together with the best fit regression line, which has the equation

$$\log_{10}(\text{reverse cumulative frequency}) = -1.757 \cdot \log_{10}(\text{bin midpoint}) + 9.727.$$

TABLE 3. Populations of the 127 largest Massachusetts cities and towns, 2010 census.

PopulationRange	Frequency	Reverse Cumulative Frequency
$< 40,000$	91	$91+25+8+1+1+1=127$
$40,000-80,000$	25	$25+8+1+1+1=36$
$80,000-120,000$	8	$8+1+1+1=11$
$120,000-160,000$	1	$1+1+1=3$
$160,000-200,000$	1	$1+1=2$
$200,000-620,000$	1	1

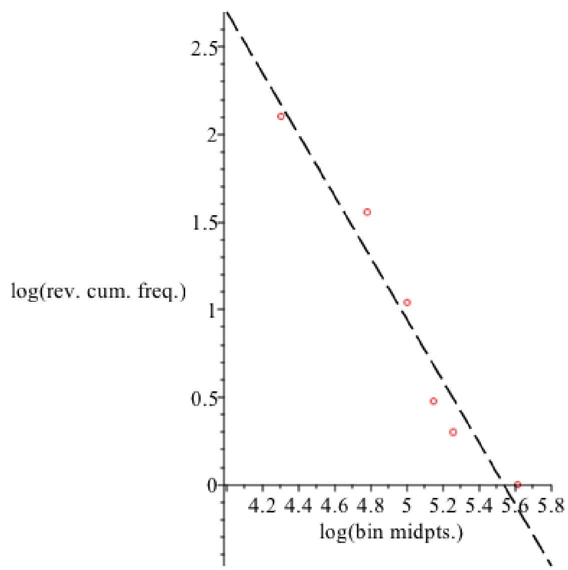


FIGURE 6.9. Fitting a power law model to the Massachusetts city and town population data.

The regression R^2 is .9497, which indicates a reasonably good fit. We can say these populations apparently follow a power law model with scaling exponent $\alpha \doteq -1.757$. However, caution must be exercised in using this approach. For instance, the estimates of the scaling exponent definitely depend on the choice of bins in constructing the histogram of the original data. For instance, note that the final bin we used above is much wider than the others–this is permissible. The alternative would be to continue with bins of width $40,000$ containing no data points until a bin containing the one largest value was found. In Exercise 10, you will redo the calculation with that choice of bins and you will find that the reported fit is essentially almost as good, but the estimated α, β parameters for the power law are different. $\triangle$

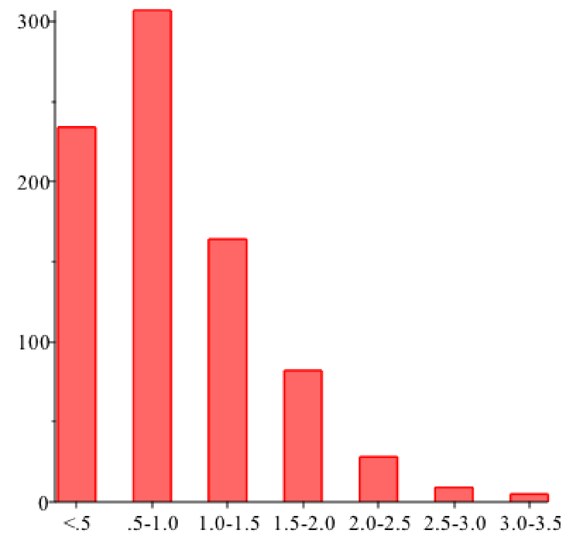


FIGURE 6.10. Histogram of earthquakes by magnitude (bins of width .5).

Here is another example where the results turn out somewhat differently.

EXAMPLE 6.8. The whole West Coast of the U.S. is a *very active* area seismically because of the numerous faults that run through this part of the country and the high level of volcanic activity. For instance, in the six-day period between June 29 and July 4, 2017, there were 829 earthquakes and other seismic events with epicenters in California and Nevada that were strong enough to be measured and recorded at the Southern California Earthquake Data Center.[8] All of them had magnitudes less than 3.5 – they were fortunately quite minor and none caused large property damage or loss of life. We will not show all the data points corresponding to the measured moment magnitudes (see Exercise 14 in Chapter 1), because the data set is so large(!) Instead, we will simply consider the frequency tabulations in Table 4. The histogram in Figure 6.10 shows the distribution of quakes by their magnitudes, computed with bins of width 0.5. Note the way the larger magnitude quakes tend to happen much less frequently than the very mild quakes with magnitudes < 1. But quakes with magnitude between .5 and 1 are more common than smaller ones. Since this data set, too, is generally following the "smaller values more common than larger ones" pattern, it makes sense to ask whether the California and Nevada earthquake magnitude data is following a power law distribution.

The computation of the reverse cumulative frequencies from the raw frequencies is shown in Table 4. We plot the log-log transformed bin-midpoint versus reverse cumulative frequency data set and the best-fit regression line in Figure 6.11. We see that this fit is nowhere near as good as some we have seen before. In fact the plot

[8]Data from `http://scedc.caltech.edu/recent/Quakes`, accessed July 5, 2017. Around 10 of the reported events were probably results of man-made explosions in quarries.

TABLE 4. Reverse cumulative frequencies of California and Nevada earthquakes by magnitude.

Magnitude	Frequency	Reverse Cumulative Frequency
$x < 0.5$	234	$234 + 307 + 164 + 82 + 28 + 9 + 5 = 829$
$0.5 \leq x < 1.0$	307	$307 + 164 + 82 + 28 + 9 + 5 = 595$
$1 \leq x < 1.5$	164	$164 + 82 + 28 + 9 + 5 = 288$
$1.5 \leq x < 2.0$	82	$82 + 28 + 9 + 5 = 124$
$2.0 \leq x < 2.5$	28	$28 + 9 + 5 = 42$
$2.5 \leq x < 3.0$	9	$9 + 5 = 14$
$3.0 \leq x < 3.5$	5	5

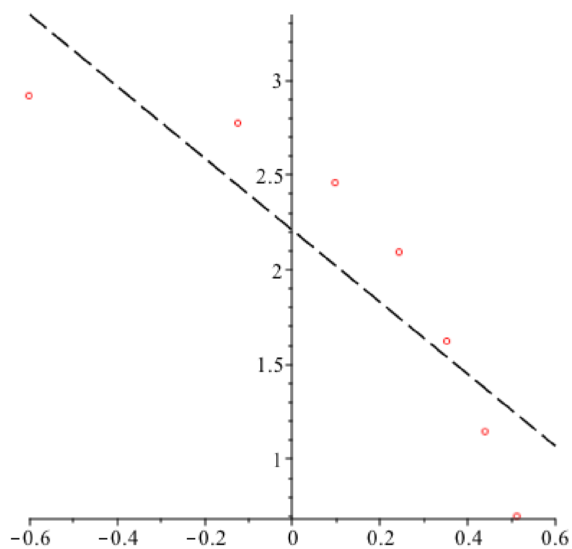


FIGURE 6.11. Log-log plot of reverse cumulative frequencies versus bin-midpoints.

definitely shows one of the warning signs that the linear regression is not fitting the data–the curvature (downward concavity) produces a definite sign pattern in the residuals. Moreover, the R^2 value, $R^2 \doteq .774$, while not terrible, is probably too far from 1 for us to draw a definite conclusion that the corresponding power law fits this distribution well. It is possible, of course, that looking at the data for a longer time span would yield different results. This 6-day period might have had uncharacteristically many quakes of magnitude beween .5 and 1 and uncharacteristically few quakes of magnitude less than .5, the apparent "culprit" for the relatively bad fit of the power law model. △

Technical note: Although the method we have outlined is simple and appealing,[9] you should be aware that it is *somewhat questionable* as a way of *inferring that a power law distribution actually fits a real world situation.* One problem is that measurement errors and random variation will be present in real-world data. Assumptions about those errors are necessary to obtain confidence bounds on the coefficients in the equation of the best-fit line (something we have not discussed yet, but an important part of using regression as a statistical tool). Even if those assumptions are true for the raw data, they are *usually not satisfied* after we apply the log-log transform. We will return to such issues later in the course. Even when a definite power law distribution should hold on theoretical grounds, the effect of random errors in the data and different binning schemes used to produce the histograms can lead to very inaccurate results for the estimated scaling exponent α and the parameter β. Hence care is required. A sounder alternative for fitting power law distributions and evaluating their goodness of fit requires more advanced mathematical and statistical techniques. If you ever need to address such questions (say in a research project later in your studies), you may want to look at the article Clauset, Shalizi, and Newman, "Power law distributions in empirical data," SIAM Review volume 51 (4) (2010) 661-703, and/or consult a practicing statistician.

6.5. Chapter Project

The chapter project for this chapter involves using the techniques we have developed to notice a pattern that holds in many types of data. The first data set you will study will be the list of populations of the 382 MSA's (Metropolitan Statistical Areas) identified by the Bureau of the Census in the United States. The MSA's are standard metropolitan areas used by government agencies and many others to study demographic and economic trends. They are designed to coincide with the major concentrations of population, *not administrative boundaries.* Thus, for instance, the MSA for Boston contains not just the City of Boston, but also the first "ring" of suburban towns around the city. The data you need is tabulated at

`https://en.wikipedia.org/wiki/List_of_Metropolitan_Statistical_Areas`

and in other places. As you can see, the first few items in the list, ranking the populations in decreasing order, are the MSA's corresponding to New York, Los Angeles, Chicago, Dallas-Fort Worth, etc. Fairly early in the list are perhaps unexpected places like Riverside-San Bernardino-Ontario, CA (located east of Los Angeles). These are not traditional "big cities," but major concentrations of population. Not surprisingly, high population states such as California, Texas, and Florida tend to have a lot of the larger MSAs. The MSA for Boston is ranked 10 on this list. Worcester, MA is included in the MSA ranked 58.

Questions. Investigate the data and try to develop answers to these questions:

(A) Does it seem as though a linear model gives a good fit between x = rank of the MSA and y = population of the MSA? Look at the value of R^2 and

[9] and analogous calculations have appeared many times in published scientific papers studying questions in biology and other areas!

the residuals for the regression – as always strong patterns in the residuals indicate a "lack of fit."

(B) What about an exponential model? Again, look at the value of R^2 and the residuals.

(C) What about a power law model? Once again, look at the value of R^2 and the residuals.

(D) It should be fairly clear from the data that the 8 or so largest MSA's are somewhat unrepresentative of the rest. What happens if you repeat parts (A), (B), and (C) on just the 9th through the 382nd?

(E) What can you say about a functional relation between $x =$ rank of the MSA, and $y =$ population based on what you have found?

(F) *Extra Credit.* Does the pattern you found here hold in other countries as well? It is fairly easy to find data online about the largest cities in countries around the world. Try the United Kingdom, India, etc. Is the pattern pretty general, or are there differences?

Now we turn to what should seem like a totally unconnected topic, namely the distribution of frequencies of words in English texts. The Corpus of Contemporary American English is a 450-million word cross-section of written and spoken English usage as it is practiced in the early 21st century. The web page

`http://www.wordfrequency.info/free.asp?s=y`

contains a listing of the top 5000 most frequently used words in contemporary English, sorted by their frequency in the Corpus. (This much is free; more detailed and more extensive lists can also be purchased!)

(G) Does it seem as though a linear model gives a good fit between $x =$ rank of the word and $y =$ frequency in the Corpus? Look at the value of R^2 and the residuals for the regression – as always strong patterns in the residuals indicate a "lack of fit."

(H) What about an exponential model? Again, look at the value of R^2 and the residuals.

(I) What about a power law model? Once again, look at the value of R^2 and the residuals.

(J) What can you say about a functional relation between $x =$ rank of a word, and $y =$ frequency in the Corpus based on what you have found?

(K) *Research Project.* Does the pattern you found here hold for other languages as well? Can you find similar compilations of words appearing in some corpus of texts for other languages? What does an analysis show?

Assignment. Write up your results and include the spreadsheets you used for these computations. As a final step before submitting your work, look up a formal statement of *"Zipf's Law."* Were your results consistent with that?

Chapter Exercises

(1) Which of the following formulas for functions can be rewritten to show that they are examples of power functions of x? For the ones that can be so rewritten, give the values of α and β from (6.1). For the ones that are not power functions, explain why they are not.

(a) $f(x) = \dfrac{5.4}{x}$

(b) $f(x) = (2.3) \cdot \sqrt[5]{x}$

(c) $f(x) = 4x^2 + 4x + 1$

(d) $f(x) = \log_{10}(4^{-(x^2)})$

(2) Find the equation of the power function $f(x) = \beta \cdot x^\alpha$ (i.e. the values of α, β) for which

(a) $f(2) = 4.5$ and $f(4) = 7.9$.

(b) $f(2.1) = 3.7$ and $f(2.4) = 3.2$.

(c) $f(4.3) = 1$ and $f(7.9) = 2$.

(3) In the previous exercise, you might have noticed that there is also a *straight line*, the graph of a linear function, passing through the points $(2, 4.5)$ and $(4, 7.9)$ from part (a) and similarly for the other parts. We saw in the text that some linear functions are also power functions. But your answers in the previous exercise are not linear functions. How can you reconcile this seeming inconsistency?

(4) Given that $f(x) = 3.4 \cdot x^{2.3}$, find the x for which $f(x) = 10$. Note: the range of a power function for $x \geq 0$ always contains all values $y > 0$ (and $y = 0$ as well, if $\alpha > 0$). So there must be such an x.

(5) Given that $f(x) = 7.8 \cdot x^{-4.3}$, find the x for which $f(x) = 10$. (See note in previous problem.)

(6) (a) Suppose the log-log plot of a data set ends up lying exactly along a line $\log_{10}(y) = 3.4 \cdot \log_{10}(x) - 7.3$. What is the relationship between y and x (without the logarithms).

(b) Same question, but the line is $\log_{10}(y) = -1.1 \cdot \log_{10}(x) + 2.7$.

(7) Enter the data from Table 1 into Excel.

(a) Verify the fit of the power law model given in Example 6.4. What is the R^2 value? Interpret the information you get from that and explain.

(b) What happens if, instead of using the midpoints of the heart rate ranges, you include two data points for each species, one for the lower boundary of the range, and one for the upper boundary of the range? How does the fitted power law model change?

(c) What happens if you fit a power law model to the data from Table 1, but you think of the heart rate as the independent variable and the body mass as the dependent variable? What power law model fits that modified data set, and how good is the fit?

(8) Enter the trout mass and length data from Table 3 from Chapter 5 into Excel and verify the calculations from Example 6.5.

(9) Using the data from Table 3 from Chapter 5:

(a) Construct a histogram for the trout masses. State clearly how you are setting up the binning scheme.

(b) Does a power law distribution apparently fit this data set? Explain.

(c) Construct a histogram for the trout lengths. Again, clearly describe your binning scheme.

(d) How well does a power law distribution apparently fit this data set? Explain.

(10) Redo the computations for fitting a power law distribution to the data on Massachusetts cities and towns from Example 6.7, but using bins of equal width equal to $40,000$. Note that this means that a number of bins with zero frequency need to be included before a bin containing the data point corresponding to Boston is reached. There will be a number of points in the reverse cumulative frequency versus bin midpoint data set where the log of the reverse cumulative frequency is equal to zero. You should see that the estimated scaling exponent α and the parameter β both change if the computation is done this way.

(11) Follow the computations for fitting a power law distribution to the data on populations of cities and towns from Example 6.7 for *your home state*. If your home state is Massachusetts, use another state(!) Use the 100 largest cities and towns. You will need to find the appropriate census data online, design a binning scheme for the histogram of populations, and then carry out the computations to see whether a power law distribution fits your data. Discuss your results.

(12) The data in Table 5 shows the 20 largest tanker oil spills in history,[10] ordered by the amount spilled in metric tons.[11] Note this does not include other oil spills due to drilling accidents, spills by oil transports other than sea-going tankers, spills by end-users, etc. Some of these spills had relatively little environmental impact because the spill occurred in open ocean, so these might not be so well-known. By the same token, some more well-known spills like the Exxon Valdez spill in Alaska in 1989 were actually farther down the list in terms of amount of oil spilled. But they had greater environmental impact because of the sensitivity of the habitats where the oil was spilled.

(a) Construct a frequency histogram for the numbers of top-20 spills by decade. When were these large spills most frequent? Why do you suppose that only one of the largest spills occurred before 1970? Does your histogram support the claim that the number of large spills has been decreasing in recent years?

(b) Does the amount spilled seem to follow a power law distribution? What are the parameters α, and β for the best-fitting model? Explain how you are drawing your conclusions.

[10] Data from `www.itopf.com`, accessed on July 7, 2017. Note: ITOPF is the International Tanker Owners Pollution Federation, a trade organization of oil tanker owners. Most of the data on this site is geared toward showing that oil spills by tankers are becoming less frequent and involve lower amounts spilled, in other words that the tanker industry's safety record is improving.

[11] One metric ton = 1000 kg, so about 2,200 lb.

TABLE 5. The 20 largest tanker oil spills in history, by amount of oil spilled.

Tanker	Year	Estimated Amount (metric tons)
ATLANTIC EMPRESS	1979	287,000
ABT SUMMER	1991	260,000
CASTILLO DE BELLVER	1983	252,000
AMOCO CADIZ	1978	223,000
HAVEN	1991	144,000
ODYSSEY	1988	132,000
TORREY CANYON	1967	119,000
SEA STAR	1972	115,000
IRENES SERENADE	1980	100,000
URQUIOLA	1976	100,000
HAWAIIAN PATRIOT	1977	95,000
INDEPENDENTA	1979	95,000
JAKOB MAERSK	1975	88,000
BRAER	1993	85,000
AEGEAN SEA	1992	74,000
SEA EMPRESS	1996	72,000
KHARK 5	1989	70,000
NOVA	1985	70,000
KATINA P	1992	67,000
PRESTIGE	2002	63,000

CHAPTER 7

Discrete-Time Dynamic Modeling and Difference Equations

7.1. Introduction

In Chapters 4, 5, and 6 we introduced techniques for modeling the relation between one independent variable and one dependent variable. In all of those examples, we started from measured values of the dependent variable for some number of input values of the independent variable and determined the models of various types that fit that measured data the best, in the sense of least squares regression. In some of the models we considered, the independent variable could be interpreted as a *time* and the results of the model could be interpreted as a description of how the dependent variable changes or evolves over time.

In this chapter, we will consider a different way to construct *dynamic models*, that is, models of changes over time. Our starting information will be different now. Instead of measured values of the dependent variable for some number of input values of the independent variable, we will start from *a statement or description of how the independent variable is changing as the independent variable changes.* This different point of view is important because many physical "laws" or scientific models can be interpreted as theoretical descriptions of how a quantity evolves or changes over time. The techniques we will introduce in this chapter will allow us to convert that sort of *dynamic description* into a list of values for the quantity that we could use to plot it as a function of time, or in some favorable cases even into a formula for the quantity as a function of time.

Instead of continuing to discuss this in the abstract, let's see a first example that relates to topics we have seen before and shows the power of approaching modeling from this new, dynamic perspective.

Example 7.1. In Proposition 5.7 we stated that any function with the property that

- *changing the independent variable by a fixed amount gives function values in a fixed proportion*

must be an exponential function. We did not justify that statement at the time. But we can now give at least a partial explanation for this claim using the new approach described above. Let us call the independent variable t (thinking of it as a time) and the dependent variable $Q(t)$. Starting from $t = 0$, consider the values $Q(0), Q(1), Q(2)$, etc. for all positive integer values of t. This is an example of a *discrete sequence* of t-values because of the discrete spacing between consecutive t-values, with spacing $\Delta t = 1 - 0 = 2 - 1 = 3 - 2 = \cdots = 1$. As a result, the values in the sequence are not "clumping closer together" or tending to a limiting value as t increases. This is the intuition behind the way the word *discreteness* is used in mathematics.

Consider the italicized statement in the bullet point at the start of this Example. If $Q(t)$ satisfies that statement, then considering the two values $Q(n+1)$ and $Q(n)$ for any n and what it means for the proportion between those two values to the the same, we see that all of the following ratios are the same:

$$\frac{Q(1)}{Q(0)} = \frac{Q(2)}{Q(1)} = \frac{Q(3)}{Q(2)} = \cdots = \frac{Q(n+1)}{Q(n)} = \cdots .$$

Let us introduce a name for this common ratio, namely $a > 0$. Then we have

$$\frac{Q(n+1)}{Q(n)} = a$$

for all $n \geq 0$, and hence

$$\boxed{Q(n+1) = a \cdot Q(n) \text{ for all } n \geq 0.} \tag{7.1}$$

This is a first example of a *difference equation* as in the title of this chapter, and we can think of it as the mathematical model equivalent to the verbal statement above.

If we look at (7.1) in the right way, we can see that it implies this formula for $Q(n)$:

$$\boxed{Q(n) = Q(0) \cdot a^n \text{ for all } n \geq 0.} \tag{7.2}$$

Specifically, this formula clearly works for $n = 0$, since $a^0 = 1$. Then, if it holds for any one integer $n = k$, that is $Q(k) = Q(0) \cdot a^k$, then (7.1) implies

$$Q(k+1) = a \cdot Q(k) = a \cdot Q(0) \cdot a^k = Q(0) \cdot a^{k+1}.$$

Hence (7.2) holds for $n = k+1$ as well. The situation is analogous to an infinite row of dominoes numbered by the integers $n \geq 0$, where, when we knock over the one for $n = 0$, that one knocks over the $n = 1$ domino, that one in turn knocks over the $n = 2$ domino, etc. Mathematically the truth of (7.2) for $n = 0$ implies the truth for $n = 1$, which then implies the truth for $n = 2$, which then implies the truth for $n = 3$, and so on for all integers $n \geq 0$.[1] Therefore $Q(t)$ is given by an exponential formula (at least for integer values of t). This gives a partial justification for what we said in Proposition 5.7. Also see Exercise 1 below, which gives an important related (and equivalent) statement. △

The point of this example is that the difference equation, the relationship in (7.1), implies that $Q(n)$ must be given by values of an exponential function. What we did in Example 7.1 can be thought of as *solving the difference equation* for the function $Q(n)$. In this chapter we will consider difference equations such as (7.1) in some generality. In some cases, we will derive techniques for solving them, that is for writing $Q(n)$ in terms of known functions. In addition, we will apply these equations to produce discrete-time dynamic models of quantities that are changing and evolving over time.[2]

---

[1]Technical note: This argument is an example of what is known as a *proof by mathematical induction*.

[2]Technical note: There is a somewhat parallel theory of *differential equations* that starts from descriptions of how functions of continuous variables are changing to derive information about the functions themselves. This is also an important modeling tool, and a core area of more advanced

7.2. Difference Equations

Let Q be a function defined for integer inputs $n \geq 0$. For us, a difference equation will be any equation relating several values of the function Q, starting with $Q(n)$ and involving some collection of the successive values $Q(n+1)$, $Q(n+2)$, $Q(n+3)$, and so forth.[3] A difference equation can also involve other known functions of the input integer value n. That equation must hold for all input values $n \geq 0$. An equation of this form is often also called a *recurrence relation* because the equation says how the values of Q for larger inputs depend on values for smaller inputs. Equation (7.1) is a first example, stating how one value $Q(n+1)$ depends on the immediately previous value $Q(n)$ and the constant a. The examples we will consider later in this chapter will come from models of some real-world situations. In those cases, the function Q will usually be specified by some sort of description, for example, as the population of a certain species of organisms in a fixed habitat as a function of time in discrete intervals of some length. But we may not know an explicit formula for it, and in fact, mathematically, we will often want to think of Q as a sort of *unknown* to be solved for from the difference equation.

If the equation involves $Q(n), Q(n+k)$ and some or all of the intermediate values $Q(n+1), Q(n+2), \ldots Q(n+k-1)$ but no other values of Q, then we say it is a kth order difference equation.[4]

Some First Order Difference Equations. Equation (7.1) is called a *first order* difference equation because the equation relates one value $Q(n)$ and the value $Q(n+1)$ which is the *first* of the following values. Another important type of first order difference equation is the so-called *affine equation:*

$$\boxed{Q(n+1) = a \cdot Q(n) + b, \text{ for all } n \geq 0.} \tag{7.3}$$

Here, a, b are constants. We will discuss modeling with difference equations of this form in the next section. Here, we will discuss how to solve an affine equation–how to produce a formula for $Q(n)$ if this difference equation holds.

Because the only difference between the affine equation in (7.3) and (7.1) is the added term b on the right, we expect that there should be a relation between the

---

applied mathematics. An introduction to differential equations is included in most Calculus 2 courses and almost all college and university mathematics major programs offer more advanced courses in differential equations as well.

[3]We will always write our difference equations this way, although in thinking about what the equation says or developing models we may also want to think of $Q(n)$ as the preceeding value to $Q(n+1)$, etc.

[4]Technical note: In case you are wondering why these are called *difference equations*, there is an explanation. Historically, difference equations were first studied as approximations to *differential* equations (see footnote 2 above), in which *derivatives* are replaced by "finite differences" such as

$$Q'(x) \doteq \frac{Q(x+h) - Q(x)}{h}$$

or

$$Q''(x) \doteq \frac{Q(x+h) - 2Q(x) + Q(x-h)}{h^2},$$

(no limits as $h \to 0$) and algebraic equations are solved to derive approximate values of Q at a discrete set of input x-values. This circumstance accounts for the way the perhaps perplexing name, "difference equation," arose historically. Example 7.9 discusses another way to understand this, at least for first order difference equations.

solutions. And indeed if we start from a known value for $Q(0)$, say $Q(0) = c$, then the affine equation implies:

$$\begin{aligned}
Q(1) &= a \cdot Q(0) + b = c \cdot a + b \\
Q(2) &= a \cdot Q(1) + b = c \cdot a^2 + b \cdot a + b \\
Q(3) &= a \cdot Q(2) + b = c \cdot a^3 + b \cdot a^2 + b \cdot a + b,
\end{aligned}$$

and so forth. Recalling that the constant $c = Q(0)$, the general pattern we see[5] here is that for any n,

$$Q(n) = Q(0) \cdot a^n + b(a^{n-1} + a^{n-2} + \cdots + a + 1).$$

From high school algebra, we recognize the sum in the parentheses (the part multiplied by the b) as the sum of a *finite geometric series*, which can also be written as

$$a^{n-1} + a^{n-2} + \cdots + a + 1 = \frac{a^n - 1}{a - 1}.$$

Hence the solution can be written as

$$Q(n) = Q(0) \cdot a^n + b \cdot \left(\frac{a^n - 1}{a - 1} \right).$$

With some more algebra, we can see an important feature of this function. Namely, if we separate the fraction $\dfrac{a^n - 1}{a - 1} = \dfrac{1}{a - 1} \cdot a^n - \dfrac{1}{a - 1}$, then we see that the first part of this can be combined with the $Q(0) \cdot a^n$ term in the formula for $Q(n)$. Doing this gives:

$$\boxed{Q(n) = \left(Q(0) + \frac{b}{a - 1} \right) \cdot a^n - \frac{b}{a - 1}.} \tag{7.4}$$

In words, the solution of the affine first order difference equation is an exponential function $c \cdot a^n$, with $c = Q(0) + \frac{b}{a-1}$, *shifted up or down* by the constant $-\dfrac{b}{a - 1}$.[6]

Example 7.2. Suppose we want to solve the affine difference equation

$$Q(n + 1) = (1.2) \cdot Q(n) + 3.4$$

with $Q(0) = 1$. Here $a = 1.2$ and $b = 3.4$. Substituting in (7.4), we see

$$Q(n) = \left(1 + \frac{3.4}{1.2 - 1} \right) \cdot (1.2)^n - \frac{3.4}{1.2 - 1} = 18 \cdot (1.2)^n - 17.$$

These values come from the function $f(x) = 18 \cdot (1.2)^x - 17$, which is related to the increasing exponential function $18 \cdot (1.2)^x$ as in Chapter 5. To get our function we subtract 17, which has the effect of shifting the whole graph down by that amount. The negative values that might be expected if you think about shifting down that far only occur for $x < 0$, though. It is clear from the difference equation that $Q(n) > 0$ for all $n \geq 0$. $\triangle$

[5]Technical note: This can be established rigorously by another proof by mathematical induction, as in Example 7.1.

[6]The shift is up if $-\frac{b}{a-1} > 0$ and down otherwise.

7.3. Some Higher Order Difference Equations (Optional)

Most of the examples we will see in the next sections involve first order difference equations. But some higher-order (i.e. second-order, third-order, etc.) differences come up in some modeling problems as well, especially in various engineering disciplines and in economics (see Exercise 10 below). So we also include a very brief introduction to some of the most important features and examples.

Example 7.3. This is an example of a *second order* difference equation:

$$Q(n+2) - n \cdot Q(n+1) + n^2 \cdot Q(n) = 1 - n, \tag{7.5}$$

because it relates one value $Q(n)$, the first succeeding value $Q(n+1)$ and the second succeeding value $Q(n+2)$. (That is, this fits the general pattern described before with $k = 2$.) Note that this relation involves several known functions of n as well, in the coefficients on the left and on the right-hand side of the equation.

A second order equation like the one in this example "kicks in" as soon as $n \geq 0$, but note that $Q(0)$ and $Q(1)$ could actually be any numbers. Knowing those *initial conditions* and the difference equation determines $Q(2)$ and then the $Q(n)$ for $n > 2$ as well. For instance if $Q(0) = 1$ and $Q(1) = 2$, then, from (7.5) with $n = 0$ we get

$$Q(2) - 0 \cdot Q(1) + 0^2 \cdot Q(0) = 1 - 0 \Rightarrow Q(2) = 1.$$

Then with $n = 1$ in 7.5,

$$Q(3) - 1 \cdot Q(2) + 1^2 \cdot Q(1) = 1 - 1 \Rightarrow Q(3) - 1 + 2 = 0, \text{ so } Q(3) = -1.$$

Then with $n = 2$ we have

$$Q(4) - 2 \cdot Q(3) + 2^2 \cdot Q(2) = 1 - 2 \Rightarrow Q(4) - 2 \cdot (-1) + 4 \cdot 1 = -1, \text{ so } Q(4) = -7.$$

Values for all $n \geq 0$ can be determined this way, one at a time. The same will be true in all of the examples we will consider. $\triangle$

It is not always easy to find an explicit formula for the function $Q(n)$ satisfying a general difference equation and some initial conditions. However, if we make some restrictions on the form of the difference equation, then this may be possible. For example there is no easy way to write down functions of n satisfying (7.5) because of the way the coefficients depend on n. One useful case where it is possible to write down formulas relatively easily is the class of *constant coefficient, linear difference equations.* For instance

$$Q(n+2) - 5 \cdot Q(n+1) + 6 \cdot Q(n) = 0 \tag{7.6}$$

is a second order, constant coefficient, linear difference equation. The "constant coefficient" descriptor means that the multiples of the various values of Q are just constants (not functions of n as in (7.5)). The term "linear" means that the $Q(n)$, $Q(n+1)$ and $Q(n+2)$ appear to the first power and not substituted into other functions or in other ways. Something like the difference equation

$$Q(n+1) = (0.5) \cdot Q(n) \cdot \left(1 - \frac{Q(n)}{40}\right)$$

is *not* linear because the way $(Q(n))^2$ would appear if we multiplied out the right hand side. We will consider some equations like this later in the chapter.

For simplicity in this elementary discussion we will restrict even farther, to *second order, constant coefficient, linear* difference equations as in (7.6). Note that the first order equation from (7.1) could be written in a very similar form:

$$Q(n+1) - a \cdot Q(n) = 0.$$

Since we know that difference equation has the exponential function solution $Q(n) = a^n$, we can "follow our noses" and see whether there are any exponential solutions of an equation like (7.6). We substitute $Q(n) = a^n$ there, then use rules for exponents and factor the result to obtain

$$a^{n+2} - 5 \cdot a^{n+1} + 6 \cdot a^n = 0 \Rightarrow (a^2 - 5 \cdot a + 6) \cdot a^n = 0.$$

This is true for all $n \geq 0$ if and only if

$$0 = a^2 - 5 \cdot a + 6 = (a-2)(a-3).$$

Hence there are two bases for exponentials that satisfy this equation, namely $a = 2$ and $a = 3$.

We claim next that we can produce many different solutions of this difference equation because:

- *For all choices of constants c_1, c_2, the function $Q(n) = c_1 \cdot 2^n + c_2 \cdot 3^n$ satisfies* (7.6).

This is easy to verify by substitution and some algebraic manipulation parallel to what we did before to find the values $a = 2, 3$. We have $Q(n+2) - 5 \cdot Q(n+1) + 6 \cdot Q(n)$

$$\begin{aligned}
&= (c_1 \cdot 2^{n+2} + c_2 \cdot 3^{n+2}) - 5 \cdot (c_1 \cdot 2^{n+1} + c_2 \cdot 3^{n+1}) + 6 \cdot (c_1 \cdot 2^n + c_2 \cdot 3^n) \\
&= c_1 \cdot 2^n (2^2 - 5 \cdot 2 + 6) + c_2 \cdot 3^n (3^2 - 5 \cdot 3 + 6) \\
&= c_1 \cdot 0 + c_2 \cdot 0 \\
&= 0.
\end{aligned}$$

This means that we can solve for c_1, c_2 to match any initial values $Q(0)$ and $Q(1)$.

EXAMPLE 7.4. Let's find a solution of $Q(n+2) - 5 \cdot Q(n+1) + 6 \cdot Q(n) = 0$ with $Q(0) = 3$ and $Q(1) = 2$. From the above, we know that all

$$Q(n) = c_1 \cdot 2^n + c_2 \cdot 3^n$$

are solutions of this difference equation. The initial conditions then say

$$\begin{aligned}
3 &= c_1 + c_2 \\
2 &= 2c_1 + 3c_2.
\end{aligned}$$

Solving simultaneously, we obtain $c_1 = 7$ and $c_2 = -4$. Hence the solution is

$$Q(n) = 7 \cdot 2^n + (-4) \cdot 3^n.$$

A function of this form is called a *linear combination* of the two exponentials 2^n and 3^n. Thinking about the range of behavior for different initial conditions, we see that because 2^n and 3^n both grow exponentially, all solutions will be unbounded except for the single solution with $c_1 = c_2 = 0$. $\triangle$

A general second order, constant coefficient, linear difference equation can be written in the form

$$Q(n+2) + r \cdot Q(n+1) + s \cdot Q(n) = f(n), \tag{7.7}$$

where r, s are constant. The equation is said to be *homogeneous* if $f(n) = 0$ for all n, and *inhomogeneous* otherwise.

We consider the homogeneous case first. We briefly describe a procedure for deriving solutions of one of these equations and given initial conditions on $Q(0)$ and $Q(1)$.

- As in the example above, the "pure exponential" solutions of the difference equation will come from values of a satisfying the quadratic equation

$$\boxed{a^2 + ra + s = 0,} \tag{7.8}$$

  which is often called the *characteristic equation* of the difference equation. So we write down this equation from the difference equation and solve it (either by factoring, or if necessary, with the *quadratic formula*).
- If there are two distinct real roots a_1, a_2, then the *general solution* has the linear combination form:

$$Q(n) = c_1 \cdot (a_1)^n + c_2 \cdot (a_2)^n.$$

  We substitute $n = 0, 1$, use the initial conditions, and solve for c_1, c_2. Apart from the numbers obtained, the process in this case is always the same as in Example 7.4.
- If there is a repeated real root a of the characteristic equation (7.8), then $Q(n) = n \cdot a^n$ is also a solution of the difference equation and the general solution has a different form, namely

$$Q(n) = c_1 \cdot a^n + c_2 \cdot n \cdot a^n.$$

  The coefficients c_1, c_2 are obtained as in the first case. See Example 7.5 below for an example.
- Finally, if the roots of the characteristic equation are not real, then they are a pair of conjugate complex numbers and the solution can be written in terms of conjugate complex exponentials. If the initial conditions are real, all the values of the function $Q(n)$ are real as well, so the complex expression can be simplified to an expression involving real exponentials and the trigonometric functions cos and sin. The algebra for doing this is somewhat beyond the scope of this elementary treatment, though, so we will not pursue this case and it will not come up in the examples we consider where we want to derive an explicit formula for the solution. However, it is the reason why particular cases of the macroeconomic model discussed in Exercise 10 exhibit *cyclical* behavior and capture an aspect of the observed business cycles in the real world(!)

Here is an example of the repeated roots case mentioned above.

Example 7.5. Consider the difference equation

$$Q(n+2) - 4 \cdot Q(n+1) + 4 \cdot Q(n) = 0,$$

with the initial conditions $Q(0) = 3$ and $Q(1) = 4$. The characteristic equation is $a^2 - 4a + 4 = (a-2)^2$, so $a = 2$ is a repeated root. $Q(n) = 2^n$ is one solution of this difference equation. But in this case we get a second solution of the form $Q(n) = n \cdot 2^n$. This is true because $Q(n+2) - 4 \cdot Q(n+1) + 4 \cdot Q(n)$ equals

$$\begin{aligned}(n+2) \cdot 2^{n+2} - 4 \cdot (n+1)2^{n+1} + 4 \cdot n2^n &= (4(n+2) - 8(n+1) + 4n) \cdot 2^n \\ &= (4n + 8 - 8n - 8 + 4n) \cdot 2^n \\ &= 0.\end{aligned}$$

Hence taking a linear combination $c_1 \cdot 2^n + c_2 \cdot n \cdot 2^n$, we have the initial conditions $Q(0) = 3$ and $Q(1) = 4$ if $c_1 + 0 = 3$, so $c_1 = 3$, and then $3 \cdot 2 + 2 \cdot c_2 = 4$. This shows $c_2 = -1$ and the solution is

$$Q(n) = 3 \cdot 2^n - n \cdot 2^n = (3 - n) \cdot 2^n.$$

Note that the step of solving for c_1 and c_2 was even easier in this example because c_2 did not appear in the equation for $n = 0$. $\triangle$

The solution of the inhomogeneous cases of (7.7) is based on the observation that *any two solutions differ by a solution of the associated homogeneous equation*, where the right-hand side $f(n)$ is replaced by 0. You will show this in Exercise 7. Hence if we can find one *particular solution* of the equation with right-hand side $f(n)$, all other solutions will equal that particular solution, plus some solution of the associated homogeneous equation. This reduces the problem to finding a particular solution. One basic technique[7] makes use of different trial solutions that contain undetermined constants based on the form of the $f(n)$. Substituting into the difference equation, a system of equations is obtained and the particular solution comes from solving them for the undetermined coefficients. We will consider some representative examples using polynomial and exponential functions $f(n)$.

EXAMPLE 7.6. Suppose we want to solve

$$Q(n+2) - 5 \cdot Q(n+1) + 6 \cdot Q(n) = n^2 + 3n + 1,$$

where the associated homogeneous equation is the one we solved in Example 7.4 above. Here $f(n) = n^2 + 3n + 1$ is a polynomial function of n of degree 2. Thinking about what sorts of functions might appear in the left-hand side to produce this $f(n)$, we see that one possibility is

$$Q(n) = An^2 + Bn + C,$$

if the undetermined coefficients A, B, C are exactly right to make

$$A(n+2)^2 + B(n+2) + C - 5 \cdot (A(n+1)^2 + B(n+1) + C) + 6 \cdot (An^2 + Bn + C)$$

equal to $n^2 + 3n + 1$. Expanding out and collecting powers of n, we have

$$2An^2 + (-6A + 2B)n - A - 3B + 2C.$$

Hence, to match $n^2 + 3n + 1$ for all n, we want

$$\begin{aligned} 2A &= 1 \\ -6A + 2B &= 3 \\ -A - 3B + 2C &= 1. \end{aligned}$$

This implies $A = \frac{1}{2}$, $B = 3$, $C = \frac{21}{4}$. The general solution will be

$$Q(n) = c_1 \cdot 2^n + c_2 \cdot 3^n + \frac{1}{2}n^2 + 3n + \frac{21}{4}.$$

We still have the constants c_1, c_2 free here so we can match any given initial values for $Q(0)$ and $Q(1)$. $\triangle$

---

[7]Technical note: This called the method of *undetermined coefficients* in more complete treatments of difference equations.

A similar computation would give a particular solution in all cases where $f(n)$ is a polynomial function. We just make a trial polynomial solution of the same degree as $f(n)$, substitute into the difference equation, and solve the resulting equations for the undetermined coefficients.

If $f(n)$ is an exponential function $f(n) = c^n$, then the trial particular solution would look like $Q(n) = A \cdot c^n$, unless c^n was also a solution of the associated homogeneous equation. If that is the case then $Q(n) = A \cdot n \cdot c^n$ would be used, unless the characteristic equation of the difference equation had the same c as a repeated root. In that last (rather uncommon) situation, the trial solution $Q(n) = A \cdot n^2 \cdot c^n$ would be used instead.[8]

Example 7.7. Suppose we want to solve

$$Q(n+2) - 4 \cdot Q(n+1) + 4 \cdot Q(n) = 2^n.$$

The associated homogeneous equation is the one we solved in Example 7.5 above, and it is one where the characteristic equation has $a = 2$ as a double root. This means that both 2^n and $n \cdot 2^n$ solve the associated homogeneous equation. Since $f(n) = 2^n$ is this same exponential function, we need a trial solution of the form $Q(n) = A \cdot n^2 \cdot 2^n$. Substituting into the difference equation and simplifying, we get

$$A(n+2)^2 2^{n+2} - 4 \cdot A \cdot (n+1)^2 \cdot 2^{n+1} + 4 \cdot A \cdot 2^n = 8 \cdot A \cdot 2^n.$$

To match $f(n) = 2^n$, we must have $A = 1/8$, and the general solution is

$$Q(n) = c_1 \cdot 2^n + c_2 \cdot n \cdot 2^n + \frac{1}{8} \cdot 2^n = (c_1 + \frac{1}{8} + c_2 \cdot n) \cdot 2^n.$$

Again, the c_1, c_2 would allow us to match any given initial values for $Q(0)$ and $Q(1)$. $\triangle$

7.4. Modeling with Difference Equations

The first example we will consider will be parallel to topics discussed in Chapter 5 on modeling with exponential functions. The idea is that knowing, for instance, that some quantity is undergoing a constant percentage increase per unit time (at least approximately) is enough to say that an exponential model is suitable (even without fitting an exponential model as we did in Chapter 5). See Exercise 1 below. Then knowing the solution of $Q(n+1) = a \cdot Q(n)$ from (7.2) lets us write down the model immediately.

Example 7.8. Over the past 50 years, a major change in the practice of mathematics and its applications studying the real world has been an amazing increase in the prevalence and power of computing resources. This has allowed the development of much more detailed and realistic models of various processes and systems. A major contributor to that increase has been the technological progress that has allowed huge numbers of electronic components (e.g. transistors) to be included in the silicon microchips making up CPU's and other computer hardware. This has been done by reducing the sizes of such components to the extent that as of

[8] Technical note: The guiding principle here is that we take $A \cdot n^k \cdot c^n$ where k is the *smallest non-negative power such that this trial solution is not a solution of the associated homogeneous equation.*

2017 the next generation of chips will have individual features about 10 *nanometers* ($\doteq 10^{-8}$ meters) in size. In 1965, Gordon Moore, an Intel engineer, predicted that *the number of transistors that could be fit on a chip would double roughly every two years.*[9]

A doubling time of 2 years means the number of transistors per chip is growing exponentially. Moreover, as we know from Exercise 6 in Chapter 5, a doubling time of 2 years corresponds to $a = 2^{1/2} \doteq 1.414$ in our general formula for exponential functions, or a 41.4% increase per year. Hence Moore's statement is equivalent to the difference equation

$$Q(n+1) = (1.414) \cdot Q(n),$$

and we know the corresponding exponential model is

$$Q(n) = Q(0) \cdot (1.414)^n, \tag{7.9}$$

where n is the number of years after $n = 0$, say the year 1965. At least through 2016 or so, a fitted exponential model to the actual data of the maximum number of transistors on the densest CPU chips matches this closely. See Exercise 11 for the details. △

We now study some examples using the first order affine difference equations from Example 7.2.

EXAMPLE 7.9. Suppose that the population of a city is growing at a net rate of 0.4% per year as a result of births and deaths, but that in addition there is a net out-migration of 500 people per year.[10] This might be happening, for instance if current residents leave to pursue better job opportunities elsewhere. If the initial population is $Q(0) = 70,000$, we might ask, assuming that the net growth rate and the number of people who emigrate every year stays the same, what will happen to the population "in the long run" – that is as t increases to larger and larger values? To answer this, we will derive a difference equation model for this situation and study the solution.

The net growth rate of 0.4% per year says that in the year between $t = n$ and $t = n + 1$,

$$\text{births} \ - \ \text{deaths} = (0.004) \cdot Q(n).$$

But during the same year, 500 of the total population are also leaving. This means that the total change from year $t = n$ to $t = n + 1$ is

$$Q(n+1) - Q(n) = (0.004) \cdot Q(n) - 500.$$

Hence, if we add $Q(n)$ to both sides and rearrange on the right we get the affine difference equation for the city population:

$$Q(n+1) = (1.004) \cdot Q(n) - 500. \tag{7.10}$$

[9]This statement is often called *"Moore's Law,"* but it is really an empirical observation and a target for the semiconductor industry to aim for, not a physical law. As of March 2017, many experts were predicting that this rate of progress in computer technology could not continue any longer, but in June 2017 IBM announced a breakthrough in silicon nanosheet fabrication that could allow for continued improvements.

[10]This means that 500 more people are moving out than are moving in, or

$$\text{(out-migration)} - \text{(in-migration)} = 500.$$

What we just did here gives another way to think about the name "difference equation," by the way(!) A first order difference equation can always be rearranged to the form

$$Q(n+1) - Q(n) = \text{some function of } Q(n) \text{ and } n,$$

and then the right-hand side gives the *difference* between $Q(n+1)$ and $Q(n)$, in terms of n and the value $Q(n)$. If we start from this form, we can always go back as well. This way of thinking is often used to derive difference equation models as we just did in this example.

We next apply the equation for the general solution of affine first order difference equations found in (7.4). With $a = 1.004$ and $b = -500$, we obtain

$$Q(n) = \left(70,000 + \frac{-500}{.004}\right) \cdot (1.004)^n + \frac{500}{.004} = 125,000 - 55,000 \cdot (1.004)^n.$$

In this example, note that $(1.004)^n$ is an increasing exponential function of n. But the multiple $55,000 \cdot (1.004)^n$ is subtracted from the 125,000, so $Q(n)$ is *decreasing* as n increases. In real-world terms, this would mean that the out-migration of 500 people per year is *not being replaced* by the natural increase of the population due to the excess of births over deaths. For example, in the first year, there were $(70,000) \cdot (0.004) = 280$ more births than deaths, but 500 people also left!

What will happen in the long run? Solving an exponential equation using logarithms as in Chapter 1, we can say that

$$125,000 - 55,000 \cdot (1.004)^n = 0$$

when

$$(1.004)^n = \frac{125,000}{55,000} \doteq 2.27 \Rightarrow n = \frac{\log_{10}(2.27)}{\log_{10}(1.004)} \doteq 205.4$$

years. Hence our model predicts that the city's population will be reduced to *zero* in about 205 years. After that point, the model gives unrealistic, negative values. Of course, it is also reasonable to ask whether the 0.4% net growth rate and the 500 person out-migration per year will persist for that long a period of time. But in any case, our model (7.10) implies that *if nothing changes*, then the city will eventually disappear, something that is not so far-fetched in many parts of the emptied-out "farm belt" of the upper mid-western U.S. at present. $\triangle$

In the previous example, the negative contribution to the difference between $Q(n+1)$ and $Q(n)$ was larger than the positive contribution. In the following example, the reverse is true.

Example 7.10. Suppose that a small glacier is losing volume due to melting of the ice at a constant rate of 2% per year, but new ice from snowfall is being laid down into the glacier each winter at the rate of .05 cubic kilometer per year. If the original volume was $Q(0) = 10$ cubic kilometers, what will happen to the glacier – will it eventually disappear, or will its volume *stabilize* at some positive value?

Here the appropriate difference equation model can again be written down by considering the difference between the end of each year and the next. Letting $Q(n)$ be the volume of the glacier at the end of year n in cubic kilometers,

$$Q(n+1) - Q(n) = -(0.02) \cdot Q(n) + .05,$$

or after rearranging

$$Q(n+1) = (0.98) \cdot Q(n) + .05 \quad \text{with} \quad Q(0) = 10.$$

This is also a first order affine difference equation with $a = 0.98$ and $b = .05$.

Using (7.4) again, we find

$$Q(n) = \left(10 + \frac{.05}{-0.02}\right) \cdot (0.98)^n - \frac{.05}{-0.02} = 2.5 + 7.5 \cdot (0.98)^n.$$

Here, $0.98 < 1$, so $(0.98)^n$ is a decaying exponential which tends to zero as n increases. Hence $Q(n)$ is decreasing with n and tending to the value 2.5 as n increases. Note that this means that the glacier *will not disappear entirely.* According to this model, it will stabilize at a volume of about 2.5 cubic kilometers. If the melting rate per year was greater (that is, $1 - a$ was smaller) or the replacement rate b from new snowfall was smaller, the value where the volume stabilized would be smaller, but it would always equal $-\frac{b}{a-1}$ according to this model. "According to this model" means, in particular, that we are assuming the melt amount per year stays constant at 2% of whatever the glacier volume is for all time in the future. That is, the model essentially includes the assumption that no continuing changes in the overall climate are going on that might increase (or decrease) the melting rate. *Models can only be as good as the assumptions that go into them!* $\triangle$

7.5. Equilibria and Stability of Solutions

In the examples so far, we have not studied how the solutions of difference equations might depend on the initial conditions and how the behavior of the solutions might change for different $Q(0)$. Part of the reason is that for the equations (7.1), there is not too much to say(!) The functions $Q(n) = Q(0) \cdot a^n$ for any one value of a are all proportional to each other, so for $Q(0) > 0$, they all either increase without bound ($a > 1$) or decrease to zero ($0 < a < 1$).

However, the affine first order difference equations from Examples 7.10 and 7.10 show that there is much more going on and more interesting behavior for solutions of other difference equations. Nevertheless, we can see what would happen for different values of $Q(0)$ by plotting solutions and/or examining the difference equation and the formulas for the solutions we derived above.

Example 7.11. Let us reconsider the difference equation from the model in Example 7.10, where we studied changes in a city's population including changes due to births, deaths, and in- and out-migration:

$$Q(n + 1) = (1.004) \cdot Q(n) - 500.$$

We saw the general solution

$$Q(n) = \left(Q(0) + \frac{-500}{.004}\right) \cdot (1.004)^n + \frac{500}{.004} = 125,000 + (Q(0) - 125,000) \cdot (1.004)^n.$$

When we look at the solution this way, we can see that there are three cases:

- If $Q(0)$ is any value greater than $125,000$, then the coefficient of the $(1.004)^n$ is positive, and the function $Q(n)$ *increases without bound.*
- If $Q(0) = 125,000$, then the coefficient of the $(1.004)^n$ is zero, and the function $Q(n)$ is *constant* with value $125,000$.
- If $Q(0)$ is any value less than $125,000$, then the coefficient of the $(1.004)^n$ is negative, and the function $Q(n)$ *decreases* and reaches values < 0 eventually.

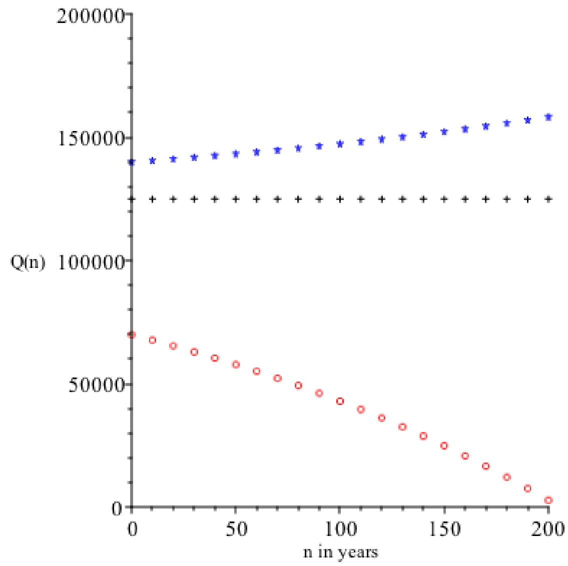


FIGURE 7.1. $Q(n) = 125,000 + (Q(0) - 125,000) \cdot (1.004)^n$, $Q(0) = 70,000$ in red, $Q(0) = 125,000$ in black, $Q(0) = 140,000$ in blue.

In the first case, we can say that with a large enough initial population, the 0.4% net growth per year is enough to replace the 500 net out-migration per year. In the third case, including the example calculations we did before, the growth is not enough to replace the losses due to out-migration and the population decreases to zero after some number of years. (After that point, the function would give negative values, and as we said before, it ceases being a realistic model for a population!) In the remaining, "middle," case the net growth exactly matches the out-migration, and the population stays constant. In other words, $Q(0) = 125,000$ is what is called a *threshold value*; only values of $Q(0) \geq 125,000$ give solutions where the population persists for all $n > 0$.

Another way to understand this pattern is to plot the function $Q(n)$ for various values of $Q(0)$, as in Figure 7.1. Note: all three solutions are plotted out to $n = 200$, but only every 10th year is shown for legibility.

The fact that $Q(n) = 125,000$ is a constant solution of the difference equation can also be seen by considering the rearranged form:

$$Q(n+1) - Q(n) = (0.004) \cdot Q(n) - 500.$$

If $Q(n)$ makes the right-hand side of this form equal to zero, then there is no change from $Q(n)$ to $Q(n+1)$ and the solution will be constant. We see

$$(0.004) \cdot Q(n) - 500 = 0 \Leftrightarrow Q(n) = \frac{500}{0.004} = 125,000.$$

So this is the constant value. $\triangle$

EXAMPLE 7.12. Next we reconsider the difference equation $Q(n+1) = (0.98) \cdot Q(n) + .05$ from Example 7.10, whose general solution is

$$Q(n) = (Q(0) - 7.5) \cdot (0.98)^n + 7.5.$$

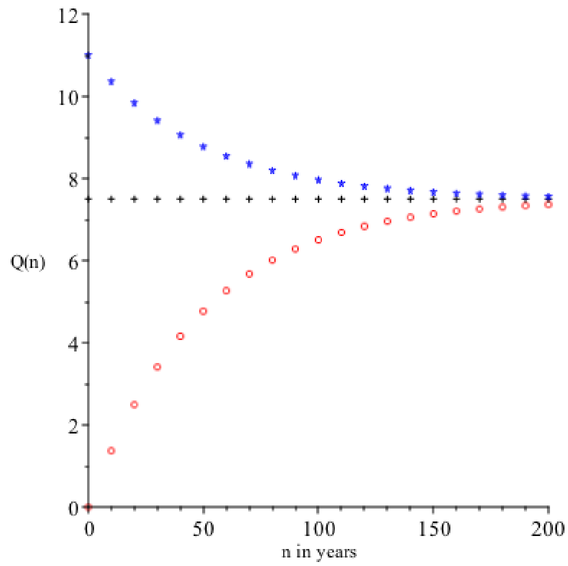


FIGURE 7.2. $Q(n) = (Q(0) - 7.5) \cdot (0.98)^n + 7.5$, $Q(0) = 0$ in red, $Q(0) = 7.5$ in black, $Q(0) = 11$ in blue.

These solutions are different from those in the previous example. There is a constant solution $Q(n) = 7.5$ for all $n \geq 0$. But since $(0.98) < 1$, the solutions for all $Q(0) > 0$ tend toward 7.5 as n increases. Three solutions are shown in Figure 7.2. Again, all three solutions are plotted out to $n = 200$, but only every 10th year is shown for legibility. $\triangle$

These examples illustrate two important concepts regarding the solutions of difference equations.

- A constant solution, $Q(n) = c$ for all $n \geq 0$, of a difference equation is also known as an *equilibrium* solution and the number c is called the *equilibrium value*. If $Q(0)$ is an equilibrium value, the corresponding solution is an equilibrium solution. For instance, $Q(n) = 125,000$ is an equilibrium solution in Example 7.11, and $Q(n) = 7.5$ is an equilibrium solution in Example 7.12.
- An equilibrium solution is said to be *stable* if all solutions that start close to the equilibrium stay close to the equilibrium (or perhaps even tend toward the equilibrium value) as n increases without bound. An equilibrium solution is said to be *unstable* if at least some of the solutions that start close to the equilibrium tend away as n increases. For example, the equilibrium $Q(n) = 125,000$ in Example 7.11 is *unstable*; the equilibrium $Q(n) = 7.5$ in Example 7.12 is *stable*.

In real-world terms, it is usually difficult to observe an unstable equilibrium with measured data. A small deviation from the equilibrium value in $Q(0)$ can and usually will produce solutions that tend away from the equilibrium value. On the other hand, stable equilibrium values will be easy to observe since even if $Q(0)$ is not exactly equal to the equilibrium value, the solution may tend to the equilibrium over time. A physical analogy may be helpful in understanding this distinction. Consider a weight on the end of a rigid rod with one end fixed in position, like the pendulum in a grandfather clock. There are two equilibrium positions: One with the weight hanging straight down; one with the weight exactly balanced at the top and the rod pointing straight up. In theory, the rod can stay motionless in either of these two positions. The first of these is stable because pulling the weight to either side by a small amount will yield motions that stay close to the equilibrium position. The second of these is unstable because moving the weight even slightly from the straight up position will cause it to move far away from that position.

To determine the equilibrium solutions of a difference equation (if there are any), the most direct way is to rewrite the equation in the form

$$Q(n+1) - Q(n) = \text{some function of } Q(n) \text{ and } n$$

discussed in Example 7.10. If there is a constant value of $Q(n)$ that makes the right-hand side *equal to zero for all* n, then that value gives an equilibrium solution.

EXAMPLE 7.13. The affine difference equation $Q(n+1) = (1.5) \cdot Q(n) - 8$ can be rewritten as

$$Q(n+1) - Q(n) = (0.5) \cdot Q(n) - 8.$$

The right-hand side of the rearranged equation is zero for all n if $Q(n) = 16$. This is the only equilibrium solution in this case.

On the other hand, suppose we started from the difference equation $Q(n+1) = (1.5) \cdot Q(n) + 8$. Rearranging as above gives

$$Q(n+1) - Q(n) = (0.5) \cdot Q(n) + 8.$$

Here the right-hand side is zero only if $Q(n) = -16$. If negative values for Q were possible, this would be a relevant constant solution. Otherwise if $Q > 0$ is necessary, we might say Q has no (positive) equilibrium values. $\triangle$

Our final example previews the difference equations considered in the next section.

EXAMPLE 7.14. For instance, consider the following difference equation and rearrange by subtracting $Q(n)$ and then factoring the right-hand side:

$$Q(n+1) = (1.1) \cdot Q(n) - (Q(n))^2 \Leftrightarrow Q(n+1) - Q(n) = (0.1) \cdot Q(n) \cdot (1 - 10 \cdot Q(n)).$$

The right-hand side of the rearranged equation is zero for all $n \geq 0$ when $Q(n) = 0$ and also when $Q(n) = 1/10$. Hence these are both equilibrium solutions. $\triangle$

7.6. Logistic Growth Models

As we pointed out in Chapter 5, the exponential models that are solutions of the difference equations

$$Q(n+1) = a \cdot Q(n)$$

may capture the behavior of biological populations in the short run. But they necessarily cease being realistic in both the cases $a > 1$ and $0 < a < 1$. In the first case, the mathematical solutions are unbounded. Hence, if we track the predictions of an exponential model, $Q(n)$ will eventually exceed any positive value, no matter how large that value is. This is impossible because real-world habitats are limited in space and in resources to support the organisms. In the second case, the values of

$Q(n)$ will eventually decline to numbers between 0 and 1 and this is also impossible if $Q(n)$ represents a number of individuals.[11]

In this section we will present and study another class of difference equation models, the so-called *logistic* models, that are designed to allow more realistic modeling of populations in resource-limited habitats. Related differential equation models were first introduced, interestingly enough, in the modeling of *human populations* by a Belgian demographer named Pierre-François Verhulst in the 1830's. But they are applicable much more generally.

ASSUMPTIONS 7.15. The assumptions underlying a logistic model for a population of a single species are the following:

(a) The population is *unstructured by age* and the habitat is *closed* (no in- or out-migration).
(b) There is a (constant) *maximum sustainable population* for the species that is a property of the habitat being modeled. This is usually called the *carrying capacity* of the habitat in biology.[12] We will denote this positive constant by M.
(c) For all n, the percentage change of the population per unit time between $t = n$ and $t = n + 1$ is proportional to the difference between 1 and the ratio $Q(n)/M$, with a positive proportionality constant that we will denote by $r \times 100\%$.

The rationale here is that $Q(n)/M$ can be thought of as the *fraction of the carrying capacity accounted for by the population* at time n. If this is less than 1, then $1 - Q(n)/M$ represents the fraction of unused capacity or "room to grow." If $Q(n)/M > 1$, on the other hand, then the population is larger than the maximum sustainable population and $1 - Q(n)/M < 0$ means the carrying capacity is exceeded. Assumption (c) implies that if $Q(n)$ is smaller than M, then $Q(n+1)$ will be larger than $Q(n)$. If $Q(n) = M$, then the percentage growth rate is actually equal to 0, and we have an equilibrium solution as discussed in the previous section. Finally, if $Q(n)$ is larger than M, the percentage growth rate is negative and $Q(n+1)$ will be smaller. The proportionality constant r, or rather $r \times 100\%$, can be interpreted as the natural net percentage population change per unit time that the species would experience in a habitat with unlimited resources.[13]

As a difference equation, the logistic model looks like this. The assumption (c) above translates into an equation

$$\frac{Q(n+1) - Q(n)}{Q(n)} = r \cdot \left(1 - \frac{Q(n)}{M}\right), \tag{7.11}$$

where r is the proportionality constant from assumption (c) above. If $Q(n) = 0$, then the percentage change between $Q(n)$ and $Q(n+1)$ is given by $r \times 100\%$.

[11] If $Q(n)$ represents a total biomass of the organisms, then values between 0 and 1 are not necessarily a problem. However, $Q(n)$ would still decline to values small enough to be unrealistic biomass values for n sufficiently large, and this is a problem.

[12] We discussed power law models for how this maximum sustainable population might depend on the average size or body mass of the animals at the start of Chapter 6.

[13] Technical note: Think of letting $M \to \infty$, so $Q(n)/M \to 0$ for any $Q(n)$.

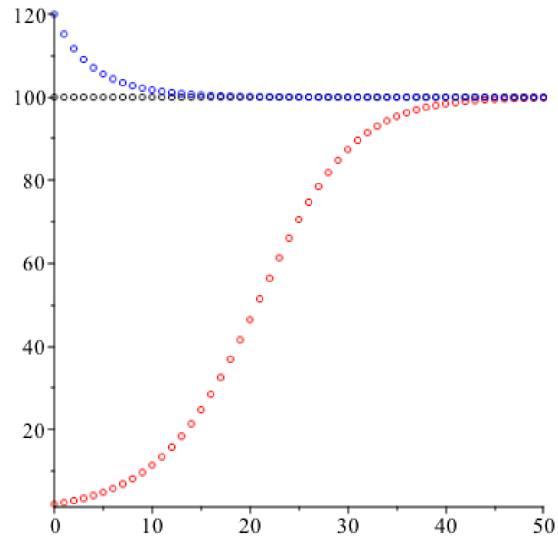


FIGURE 7.3. Several solutions of (7.13).

We will usually rearrange this to the standard *logistic equation* form:

$$\boxed{Q(n+1) = (1+r)\cdot Q(n) - \frac{r\cdot (Q(n))^2}{M}.} \tag{7.12}$$

EXAMPLE 7.16. For instance, if Q is increasing at a 20% rate per unit time when $Q = 0$, then $r = .2$. Say the carrying capacity is $M = 100$. Then the corresponding logistic difference equation as in (7.12) is

$$\begin{aligned} Q(n+1) &= (1.2)\cdot Q(n) - \frac{(0.2)\cdot (Q(n))^2}{100} \\ &= (1.2)\cdot Q(n) - (0.002)\cdot (Q(n))^2. \end{aligned} \tag{7.13}$$

Figure 7.3 plots several solutions of this difference equation, with initial conditions $Q(0)$ between 0 and 100, equal to 100 and greater than 100. In the population model context, any value of $Q(0) \geq 0$ would be a possibility. There are two equilibrium solutions of this equation, namely $Q(n) = 0$ and $Q(n) = 100$ for all n. This is easiest to see from the unsimplified form (7.11) if we multiply through by $Q(n)$. From the plots, we guess that 0 is an unstable equilibrium and 100 is a stable equilibrium. $\triangle$

The logistic difference equation (7.12) is not linear. A byproduct of this fact is that, unlike the affine first order equations considered before, logistic equations do not possess convenient closed formulas for the solutions (except for the equilbrium solutions!) This means that if we want to investigate solutions of a logistic equation, we may need to generate a list of the values $Q(n)$ one at a time for some range of n and possibly plot the data points $(n, Q(n))$. This can be done in a number of different ways either by hand or using software. Excel can be adapted for this purpose quite easily by making use of the addressing conventions when formulas are copied from one cell into another. For instance, the following commands will generate the values of n and $Q(n)$ for the solution of (7.13) and initial condition $Q(0) = 20$, $n = 0, \ldots, 20$ in two parallel columns, then generate a point plot of the $(n, Q(n))$.

- Enter 0 in cell `A1` and 20 in cell `B1`. (You might also want to enter text column headings in those columns; if so start in row 2 instead of row 1.)
- Enter the formula (macro) `=A1+1` in cell A2 and press ENTER/RETURN. Enter `=(1.2)*B1-(0.002)*B1^2` in cell `B2` and press ENTER/RETURN.
- Copy the formula from cell `A2`, highlight cells `A3` through `A21` and paste. You should see the numbers $2, 3, \ldots, 20$ in those cells. What is happening is that when you paste, the row number in the address `A1` from is increased by 1 each time it is inserted in a new row down the column. This is happening because we did *not* use `$A$1` to make it an absolute address.
- Similarly if you copy and paste the formula from `B2`, the copy in cell `B3` will have the address `B1` updated to `B2`, so the formula will be evaluated using the value in `B2`, the copy in cell `B4` will use the value in `B3` and so forth. The result is to compute each value of $Q(n+1)$ using the previous value $Q(n)$, just as we want.
- With the parallel columns `A` and `B` rows `1` through `21` (i.e. the rectangular block `A1:B21`, you can generate a point plot (scatter plot) using the Insert/Chart Design options and this shows the $(n, Q(n))$ data points.

Modeling with Logistic Equations. In applications, one common procedure is to take a real-world data set and *fit* a logistic model to it in order to estimate the proportionality constant r and/or the carrying capacity M, and then use that information and the solutions of the logistic differential equation to understand projected behavior (assuming that r and M remain constant). The quality of the results will usually depend on how well those assumptions match the real-world situation(!)

EXAMPLE 7.17. In Exercise 11 from Chapter 5, we discussed fitting linear and exponential models to the U.S. population data from 1790 through 2010. This is reproduced for the reader's convenience in Table 1. Let us now reconsider this data and see the process of fitting a logistic model. For simplicity, let us say that n here represents the number of 10-year periods after 1790. So $n = 0$ corresponds to the first census in 1790 and $n = 22$ corresponds to the most recent census in 2010.[14] The idea comes directly from (7.11), the translation of assumption (c) above describing logistic models. We need to compute the approximate percentage changes *per unit time* first (but remembering that the unit of time is 10 years, not 1 year. For this we will use the values $\frac{Q(n+1)-Q(n)}{Q(n)}$.

To make a point, we will only use the data from the censuses starting from 1790 through 1930. We fit a linear function of Q, say $\hat{m} \cdot Q + \hat{b}$, to the data points

$$\left(Q(n), \frac{Q(n+1) - Q(n)}{Q(n)}\right), \quad n = 0, 1, 2, \ldots, 14.$$

[14] I am writing this in July, 2017.

TABLE 1. U.S. Population According to Federal Census Records

Census	Population	Census	Population
1790	3,929,214	1800	5,308,483
1810	7,239,881	1820	9,638,453
1830	12,866,020	1840	17,069,453
1850	23,191,876	1860	31,443,321
1870	38,558,371	1880	50,189,209
1890	62,979,766	1900	76,212,168
1910	92,228,496	1920	106,021,537
1930	123,202,624	1940	132,164,569
1950	151,325,798	1960	179,323,175
1970	203,211,926	1980	226,545,805
1990	248,709,873	2000	281,421,906
2010	308,745,538		

Then from the right side of (7.11),

$$\hat{m} \cdot Q + \hat{b} \doteq r \cdot \left(1 - \frac{Q}{M}\right) = \frac{-r}{M} \cdot Q + r,$$

we estimate $r \doteq \hat{b}$ and

$$\hat{m} \doteq \frac{-r}{M} \Rightarrow M \doteq \frac{-r}{\hat{m}} = \frac{-\hat{b}}{\hat{m}}.$$

The results are shown in Figure 7.4. As you can see from the plot, the fit is reasonably good; the regression $R^2 \doteq .91$. But in any case, the equation of the regression line is approximately

$$0.364 - (2.208 \times 10^{-9}) \cdot Q$$

which says our estmated $r \doteq .364$ and

$$M \doteq \frac{-.364}{-2.208 \times 10^{-9}} \doteq 1.649 \times 10^{8}.$$

The plot in Figure 7.5 shows the model's predicted values for 1790 through 2010 (in black) together with the actual population data (in red). The model's predicted values are quite close until $n = 160$ (1950), and after that the model values are leveling off and approaching the predicted carrying capacity of $164,900,000$. On the other hand the actual population is still increasing rapidly.

The point we wished to make here is that even though the logistic fit is quite good up until about 1940, there are a number of features of this real-world situation that do not match the underlying assumptions of the logistic model as described before. First, the U.S. population was definitely not evolving in a closed habitat. The geographical area under the jurisdiction of the U.S. government was not even

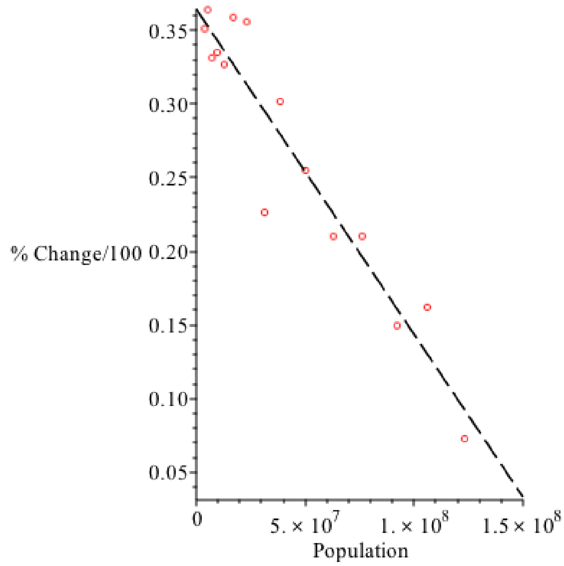


FIGURE 7.4. Fitting a logistic model to the data from 1790 - 1930.

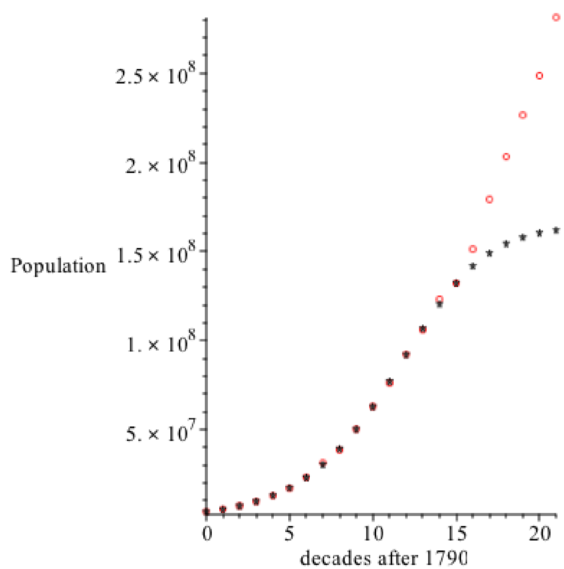


FIGURE 7.5. Logistic model from the 1790 - 1930 data (in black) and actual population (in red).

constant – it was continually expanding through the 19th century. Moreover, the population grew as a result of large numbers of immigrants from Europe, Latin America, Asia, as well as enslaved persons from Africa before 1860. Second, the assumption that the carrying capacity was a constant does not reflect the fact that improvements in all sorts of technologies have certainly made it possible to support larger human populations in the same land area. In fact, the model's estimated carrying capacity of $164,900,000$ was exceeded almost as soon as the interval we modeled was complete. Hence it is probably to be expected that the logistic model only fits the data reasonably well in the period corresponding to the data points we used. $\triangle$

Logistic models have some other unexpected and interesting properties "up their sleeves." See Exercise 17 below in particular.

7.7. Chapter Project

Environmental scientists often try to estimate populations of plant or animal species and understand to what extent they can be used as resources by humans without being depleted. For instance, populations of wild fish and other marine creatures around the world have been a major source of food for humans for many years. Yet there is evidence that many of them have been overfished and there is fear some of them may be headed toward extinction. For example, following 500 years of fishing, by the summer of 1992, the biomass of northern cod observed in Atlantic waters off the coast of Newfoundland had fallen to an estimated 1% of its previous levels. As a result, the federal government of Canada declared a moratorium on cod fishing, hoping to give the cod populations time to recover. The economic and social impacts on the human population of Newfoundland were severe, since many of the people in the area derived their entire livelihood from cod fishing and others depended on cod fishermen as customers for their businesses. An estimated $35,000$ people lost their jobs as and the whole society of Newfoundland has not really recovered to this day. As of around 2010, there were some encouraging signs that the cod fishery might be recovering, but the effects of other factors such as changes in ocean temperature and loss of populations of the food species that cod eat have kept the ultimate fate of the north Atlantic cod fishery uncertain. Similar decreases have been observed in New England cod populations more recently and the U.S. government has instituted more and more stringent fishing limits to try to avert collapses here as well. The fishing industries of Maine and Massachusetts are under similar pressures.

In this project, you will study various models of a fishery including effects from fishing by humans. Let $P(n)$ represent the total mass of mature Pacific halibut in units of 10^6 kg. We will model the wild halibut biomass *without any fishing* by the following logistic difference equation

$$P(n+1) = (1.71) \cdot P(n) - (.00875) \cdot (P(n))^2. \tag{7.14}$$

Here $r = .71$ and $r/M = .00875$.

Questions.

(A) *"Greedy" harvesting.* Suppose the halibut stock started out at 95% percent of the carrying capacity according to the model (7.14). But in one massive fishing effort, the halibut biomass is reduced all the way down to 1×10^6 kg (say all within one year). If no further fishing is allowed until stocks recover to 95% of the carrying capacity, how long will that take, according to the model? Estimate by computing a solution of the difference equation (7.14). What would be the *average fish amount taken per year* if this process of massive fishing followed by fallow time to allow recovery to 95% of the carrying capacity was done repeatedly over a long period?

(B) *Constant harvesting.* One way to make use of a resource like the halibut fishery that is less drastic than the "greedy" approach in (A) is to take some constant amount of fish every year.

(1) Suppose that everything remains as in (7.14) above, but some constant amount h (in 10^6 kg) of halibut biomass is removed each year via fishing.[15] What modified difference equation models this situation? (Think about the derivation of (7.12) and take the fishing amount h into account.)

(2) Investigate the solutions of your constant harvesting difference equation from part (1) if the fishing term is each of these values: $h = 5, 10, 14, 20$, one at a time. Choose enough different $P(0)$ values for each so that you think you see the whole picture and then describe what is happening in words. In particular, for each h how many different equilibrium solutions are there? Where are they located? How do they change as h increases? Are they stable or unstable?

(3) By rewriting your difference equation from part (1) in the form

$$P(n+1) - P(n) = \cdots ,$$

what is the *maximum value* of h for which the equation still has a stable equilibrium? (This question can be answered by means of algebra alone if you think about it the right way!)

(4) What should it mean to say that a fishing level h is *sustainable*?

What is the maximum sustainable constant fishing level? Does the answer depend on what the initial value $P(0)$ at the start of the fishing intervention is?

(5) What would be the *average fish amount taken per year* if constant harvesting at the maximum sustainable level is done repeatedly over a long period?

(C) *Proportional harvesting.* Instead of taking a constant amount of fish, we could also take a constant proportion of whatever fish biomass is present.

(1) Next, suppose that everything remains as in (7.14) above, but instead of a constant amount, suppose that a constant *proportion* p of the halibut biomass (whatever it is) is removed each year via fishing.[16] What modified difference equation models this situation? (Think about the derivation of (7.12) and take the proportion removed by fishing into account.)

(2) Investigate the solutions of your constant harvesting differential equation from part (1) if the fishing term is each of these values: $p = 0.1, 0.3, 0.5, 0.8$, one at a time. Choose enough different $P(0)$ values so that you think you see the whole picture and then describe what is happening in words. In particular, for each value of p how many

[15] Think h means a "harvesting level," hence the notation.

[16] Think $0 < p < 1$ with $p = 1$ meaning all of the fish are removed.

different equilibrium solutions are there? Where are they located? How do they change as h increases? Are they stable or unstable?

(3) By rewriting your difference equation from part (1) in the form

$$P(n+1) - P(n) = \cdots,$$

what is the p for which the halibut population starts to "crash" for all $P(0)$? (This question can be answered by means of algebra alone if you think about it the right way. And the answer should make biological sense too!)

(4) What should it mean to say that a fishing proportion p is *sustainable*? What values of p are sustainable? Does the answer depend on what the initial value $P(0)$ at the start of the fishing intervention is?

(5) What would be the *average fish amount taken per year* if proportional harvesting at the level $p = .3$ is done repeatedly over a long period?

Assignment. Compare the strategies in parts (A), (B), (C) from the point of view of their effect on the halibut fishery and the average amounts taken per year. If you were going to recommend one, which would it be? Explain how you are making your determination.

Chapter Exercises

(1) Suppose a quantity $Q(t)$ satisfies the property that for all $n \geq 0$, *the percentage change in* Q *between* $t = n$ *and* $t = n + 1$ *is always the same*, say $r\%$. Show that $Q(t)$ satisfies the difference equation

$$Q(n+1) = \left(1 + \frac{r}{100}\right) Q(n) \quad \text{for all } n \geq 0.$$

Deduce that $Q(n) = Q(0) \cdot \left(1 + \frac{r}{100}\right)^n$ for all $n \geq 0$.

(2) Continue the calculation begun after Example 7.3 to find the values $Q(5)$, $Q(6)$, and $Q(7)$ if $Q(n)$ satisfies (7.5) and $Q(0) = 1$, $Q(1) = 2$.

(3) Follow the calculation begun after Example 7.3 to find the values

$$Q(2), Q(3), Q(4), Q(5), Q(6), Q(7)$$

if $Q(n)$ satisfies (7.5), but now use different intiial conditions $Q(0) = 3$ and $Q(1) = 1$. Your values should be completely different from what you saw in the previous Exercise.

(4) Solve the following first order difference equations with the given initial conditions.
   (a) $Q(n+1) = (1.8) \cdot Q(n)$ with $Q(0) = 3.4$.
   (b) $Q(n+1) = (.78) \cdot Q(n) - .03$ with $Q(0) = 4.3$

(5) The first order difference equation $Q(n+1) = (Q(n))^2/4$ does not fall into any of the general patterns we discussed in the text.
   (a) Compute the values $Q(1), \ldots, Q(4)$ if $Q(0) = 5$.
   (b) Compute the values $Q(1), \ldots, Q(4)$ if $Q(0) = 1/2$.
   (c) By examining the values of $Q(n)$ for a general initial condition $Q(0) = a$, guess a formula for the general solution.

(d) Does this differential equation have any equilibrium solutions? If so, what are the equilibrium values? What appears to be true about their stability?

(6) (For the optional §7.3.) Consider the difference equation

$$Q(n+2) - 2 \cdot a \cdot Q(n+1) + a^2 Q(n) = 0.$$

Show that the characteristic equation has a as a double root and that every function of the form $Q(n) = c_1 \cdot a^n + c_2 \cdot n \cdot a^n$, with c_1, c_2 constant, is a solution.

(7) (For the optional §7.3.) Show that if $Q_1(n)$ and $Q_2(n)$ are any two solutions of the inhomogeneous difference equation $Q(n+2)+r \cdot Q(n+1)+s \cdot Q(n) = f(n)$, then the difference $Q(n) = Q_1(n) - Q_2(n)$ solves the associated homogeneous difference equation $Q(n+2) + r \cdot Q(n+1) + s \cdot Q(n) = 0$.

(8) (For the optional §7.3.) Solve the following difference equations and initial conditions.

(a) $Q(n+2) + 7 \cdot Q(n+1) + 6 \cdot Q(n) = 0$, $Q(0) = 4$, $Q(1) = 2$.

(b) $Q(n+2) + 6 \cdot Q(n+1) + 9 \cdot Q(n) = 0$, $Q(0) = 1$, $Q(1) = 1$.

(c) $Q(n+2) - 6 \cdot Q(n+1) + 8 \cdot Q(n) = n+3$, $Q(0) = 1$, $Q(1) = 3$

(d) $Q(n+2) + 4 \cdot Q(n+1) + 3 \cdot Q(n) = 5^n$, $Q(0) = 2$, $Q(1) = 1$.

(9) (For the optional §7.3.) Consider the inhomogeneous second order linear difference equation with constant coefficients:

$$Q(n+2) + r \cdot Q(n+1) + s \cdot Q(n) = c.$$

(a) If c is constant and $1 + r + s \leq 0$, then show that $Q(n) = \frac{c}{1+r+s}$ is an equilibrium solution.

(b) Assume the characteristic equation has two distinct real roots or a real double root. Explain why this equilibrium is *stable* only in the case that the roots of the characteristic equation are all less than 1 in absolute value.

(10) (For the optional §7.3.) In 1939, the economist Paul Samuelson (1915-2009) developed the *multiplier-accelerator* model,[17] a simple difference equation macroeconomic model. This was a major advance because in some cases (i.e. for some values of the parameters α, β in the coefficients) it gave solutions exhibiting *cyclical* behavior like the periodic boom-bust cycles that are observed in real-world economic systems. In this problem, you will work through the derivation of the multiplier-accelerator model, a second-order linear difference equation. The cyclical behavior occurs when the characteristic equation has complex conjugate roots as discussed in the text. The function $Q(n)$ in the model represents the total level of activity at time $t = n$ in the economy of a whole nation. By assumption, this exactly matches the spending intentions of the government, the consumption intentions $C(n)$ of households, and the investment intentions $I(n)$ of businesses. For simplicity, say the government spending is constant $= 1$, then

$$Q(n) = 1 + C(n) + I(n). \tag{7.15}$$

[17] "Interactions Between the Multiplier Analysis and the Principle of Acceleration," Review of Economic Statistics. 21 (2): 75–78.

(a) The "multiplier effect" is the statement that $C(n)$ is proportional to the economic activity, but with a time lag: $C(n) = \alpha \cdot Q(n-1)$ for some constant α. The "accelerator effect" is the statement that the business investment reacts to how consumption is *changing*: $I(n) = \beta \cdot (C(n) - C(n-1))$, for some constant β. Substitute these into (7.15) to obtain the *multiplier-accelerator model*:

$$Q(n) = 1 + \alpha \cdot (1+\beta) \cdot Q(n-1) - \alpha \cdot \beta \cdot Q(n-2).$$

Reindex and rearrange algebraically to put this in our standard form:

$$Q(n+2) - \alpha(1+\beta) \cdot Q(n+1) + \alpha \cdot \beta \cdot Q(n) = 1.$$

(b) In the rearranged form in part (a), suppose $\alpha = 0.4$ and $\beta = 2$. Show that the characteristic equation of the difference equation has no real roots.

(c) Continuing with the values of α, β in part (b), and initial conditions $Q(0) = 1$, $Q(1) = 1.1$ compute the values of $Q(n)$ for $n = 2, \ldots, 20$. (You may want to use Excel for this.) Describe what the solution is doing in words.

(11) Table 2 gives the number of transistors on some CPU chips as a function of time. Fit an exponential model to this data and compare with (7.9) in Example 7.8

TABLE 2. "Moore's Law"

CPU	Year	Number of Transistors
4004	1971	2250
8008	1972	2500
8080	1974	5000
8086	1978	$29,000$
286	1982	$120,000$
386	1985	$275,000$
486 DX	1989	$1,180,000$
Pentium	1993	$13,100,000$
Pentium II	1997	$7,500,000$
Pentium III	1999	$24,000,000$
Pentium 4	2000	$42,000,000$
Xeon	2008	$1,900,000,000$

in the text.

(12) Suppose that in 2017, the population of a country was $34,000,000$. The death rate exceeds the birth rate by 0.3% per year, but each year $100,000$ more people are entering the country from the outside than leaving.

(a) Express the information above as a difference equation model.

(b) Using your model, compute the projected population each year until the year 2037. Describe what is happening in words.

(c) What is the equilibrium level of your difference equation model?

(13) For each of the following logistic models and initial conditions,
(i) Identify the r, M parameters from the model,
(ii) Plot the solution using Excel for $n = 1, \ldots, 50$,
(iii) Discuss the results.

(a) $Q(n+1) = (1.03) \cdot Q(n) - (.006) \cdot (Q(n))^2$, $Q(0) = 0.8$.

(b) $Q(n+1) = (1.34) \cdot Q(n) - (.0009) \cdot (Q(n))^2$, $Q(0) = 420$.

(c) $Q(n+1) = (1.86) \cdot Q(n) - (.0048) \cdot (Q(n))^2$, $Q(0) = 1.3$.

(14) Many biologists are not convinced that logistic models are the best choices for population modeling in limited environments and a number of other sorts of models have been proposed as substitutes.[18] Some of the objections involve point (a) in Assumptions 7.15 – that there is no age structure in the population. In most animal species for instance, immature juvenile individuals cannot reproduce and adults can live long enough that they can no longer reproduce. More realistic models would take that age stratification into account by subdividing the population into several groups and tracking them separately. We will study examples of such models in the next chapter. But there are also criticisms based on point (c) in Assumptions 7.15. We consider the logistic model in the form

$$Q(n+1) - Q(n) = r \cdot Q(n) - \frac{r(Q(n))^2}{M}.$$

(a) Explain why the $\frac{r(Q(n))^2}{M}$ can be interpreted as the *number of deaths per unit time due to overcrowding, competition for resources, etc.* (i.e. excess deaths in addition to the natural deaths that would occur in an unlimited habitat).
(b) Explain why it is plausible that this number of excess deaths is proportional to $Q(n) \cdot Q(n)$.
(c) Discuss whether it is plausible that this number of excess deaths should be also proportional to the natural net growth rate r.

(15) National censuses are also taken every 10 years in the United Kingdom, in years ending in 1. Follow what we did in Example 7.17 in the text to fit a logistic model to the data at

`https://en.wikipedia.org/wiki/Demography_of_the_United_Kingdom`[19]

to fit a logistic model to the population of the United Kingdom, using the data from the table marked *Population at Census Dates*. Use data for the years 1851, 1861, 1871, 1881, 1891, 1901, 1911, 1921 (in the column marked *at start of period*). What are the estimated r, M for your model? How well does your

[18]Technical note: One such type is the class of so-called Gompertz models. These are most commonly used in the continuous time case as differential equation models, but discrete-time versions have also appeared in the literature.

[19]consulted July 12, 2017.

model fit the actual population over that time? How well does it fit in the period following 1921?

(16) The biologist G.F. Gause published a well-known book called *The Struggle for Existence* in 1934. Part this was devoted to studies of one-celled organisms such as yeasts and paramecia in limited environments demonstrating the effects of resource limitations and competition. One of his experiments yielded data similar to that in Table 3 for the combined numbers of two species of paramecia:[20]

TABLE 3. Paramecium populations

Day	Population
0	21
1	67
2	104
3	137
4	165
5	170

(a) Following what we did in Example 7.17, fit a logistic model to this data set.
(b) Use your model to predict the paramecium population out to 25 days.
(c) Gause wrote "At a certain moment the possibility of continued growth in a given microcosm is apparently exhausted, and with a continuously maintained level of nutritive resources, a certain equilibrium of population is established." Where does your model predict that occurs here?

(17) *Logistic Models and "Chaos."* The biologist Robert May pointed out in 1976[21] that the solutions of logistic models can exhibit surprisingly complicated behavior when the r parameter in (7.12) is larger than the small values we have studied in the examples in the text. Moreover, large values of r can certainly occur for species of insects, microorganisms, etc. So this complicated behavior is not only of theoretical interest; it is something that might be observed "in the wild" in some cases. In this exercise, you will investigate some of this complicated behavior. Consider the logistic equation

$$Q(n+1) = (1+r)\cdot Q(n) - r\cdot (Q(n))^2,$$

where we have taken $M = 1$ for simplicity. Using the initial condition $Q(0) = 0.3$, investigate the following cases by computing and plotting solutions using Excel. You will probably want to continue until $n = 100$ or so in all cases to see what is going on.

[20] These numbers are somewhat different than his from for the purposes of this problem. He also continued the experiment to 25 days and eventually observed oscillations about an equilibrium value.

[21] "Simple mathematical models with very complicated dynamics" Nature 261(5560):459–467.

(a) Take $r = 2$ first. What do the solution values do as n increases?

(b) Now repeat for $r = 2.5$. Look at the computed values carefully in addition to the plot. What is happening here?

(c) Now repeat for $r = 2.5441$. Again look at the computed values carefully.

(d) Repeat for $r = 2.7$.

The reasons for this cascade of more and more complicated behavior and a complete description of what is happening here (which is a LOT more complicated than the small number of different cases seen above!) have been rigorously established.[22] What you are seeing here is a few steps in the so-called *period-doubling cascade* and the solutions for $r \geq 2.7$ are said to exhibit *chaotic dynamics*. Based on what you saw in part (d), that should seem like a reasonably descriptive name!

[22] i.e. with full mathematical proofs of precise statements of what is going on

CHAPTER 8

Modeling With Systems of Difference Equations

8.1. Introduction

In our discussion of the assumptions behind logistic models for population dynamics in Chapter 7, we mentioned that the assumption that all individuals in the population are essentially treated as interchangeable is quite unrealistic from the biological viewpoint. Animal species (for which sexual reproduction is the most common strategy) have populations made up of males and females. Logistic models do not take that into account. Moreover, recall that we said in Assumptions 7.15 that the population was "unstructured by age." In particular, all live individuals are treated as being able to reproduce. Thus, there is no juvenile period where individuals are not yet sexually mature. And similarly, there is no senescence period where individuals have aged out of the ability to reproduce. This is not representative of the facts of life for many longer-lived animal species. For instance humans do not become sexually mature until they reach puberty.[1] At the other end of the life cycle, human females typically live another 20 to 30 years after they reach menopause, when they become infertile.[2] Human males typically remain able to reproduce somewhat longer.

If we consider what it would take to include features like a juvenile period or a senescence period in a population model, then it is clear that the total population $P(n)$ at time n should be stratified or structured into (at least) three groups, the juveniles J, the mature individuals M and the senescent individuals S with

$$P(n) = J(n) + M(n) + S(n).$$

In addition we might want to subdivide J, M, S even further into the males and females within each age cohort if there were significant differences between the two sexes. This would require something like six dependent variables $J_m(n)$ for the male juveniles, $J_f(n)$ for the female juveniles, etc. A population model would then describe how each of these contribute to the corresponding $J_m(n+1)$, $J_f(n+1)$, $M_m(n+1)$, $M_f(n+1)$, $S_m(n+1)$, and $S_f(n+1)$ at the next time step. Hence what we will be describing is a *simultaneous system* of several first order difference

[1] There has been a decrease in the average age of onset of puberty over time. For instance, a study reported that the average age of onset of puberty in girls in the Germany was 16.6 years in 1860, but this had fallen to 10.5 years in 2010. Similar decreases have been observed in many other countries as well, including the U.S. A parallel trend for boys has been observed, but with ages about one or two years greater than for girls. There is no clear consensus about the cause, although improvements in nutrition and greater prevalence of certain chemicals in the environment tied to production of sex hormones have been suggested. See `https://www.theguardian.com/society/2012/oct/21/puberty-adolescence-childhood-onset`, accessed July 18, 2017.

[2] The average age of onset of menopause is around 51 years for U.S. women.

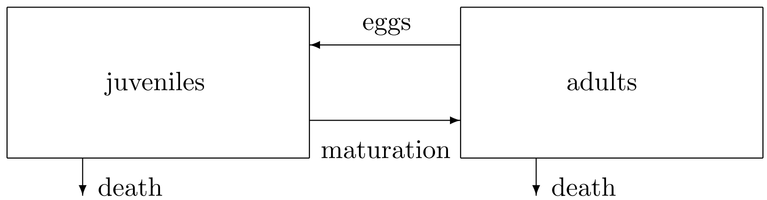


FIGURE 8.1. A schematic two-compartment population model.

equations for the age and gender strata of the population as functions of the discrete time variable.

In this chapter, we will study such structured population models in detail first, then look at several other types of models given by systems of difference equations. The models we will develop are also commonly called *compartment models* since developing them will often involve analyzing how each of the different strata or compartments at time n contributes to the corresponding compartments at time $n + 1$.

8.2. Age-structured Population Models

Let's begin with a very simple example that shows how to set up one of these compartment models.

EXAMPLE 8.1. Instead of a long-lived animal species, let us consider a situation more typical of *insects* where the population is made up of only *juveniles* and *adults*. We will also treat only the *female* members of the population. We state our assumptions in words first, then consider how to translate them into a more formal mathematical model. We assume:

- Juveniles enter the juvenile population by way of hatching of eggs laid by the adults in the previous time step
- Juveniles leave the juvenile population either by becoming adults, or by dying
- Adults enter the adult population from the juvenile population when they mature
- Adults leave the adult population by dying

At this level of detail we are not distinguishing between the possible causes of death. This could be due to predation, to "natural causes," to diseases, or something else.

It will be helpful to visualize all of this by drawing the following sort of *compartment diagram*, where the boxes marked J and A represent the juvenile and the adult strata of the insect population, respectively.[3]

To turn Figure 8.1 into a model, we need to make some assumptions.

- Let us say that the deaths from the juvenile group make up some constant fraction $0 < d_J < 1$ of whatever that population is. That is, the constant

[3]You may notice a similarity between this sort of diagram and the energy flow diagrams we studied in the Part I Summary Project from Chapter 3. Those were slightly different though, in that they did not contain the *dynamic* aspect of a change from one time step to the next.

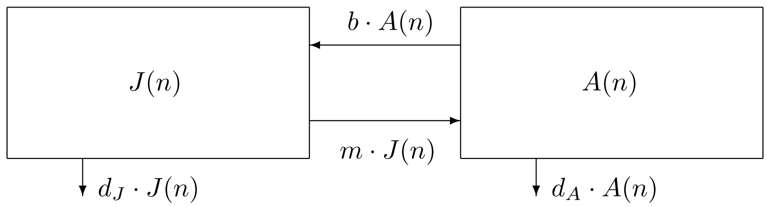


FIGURE 8.2. A simple two-compartment population model.

d_J represents the *juvenile death rate per unit time*. Therefore, the arrow marked death leading out of the J box in Figure 8.1 will represent a loss of $d_J \cdot J(n)$ individuals.

- Similarly say the deaths from the adult group represent a loss of $d_A A(n)$ where d_A is a constant adult death rate per unit time. Note: The juvenile death rate and the adult death rate could conceivably be different[4] so we have given them different names to allow for this.
- The arrow marked "maturation" represents the movement of (former) juveniles into the adult group when they mature. Let us say that this proportion is some constant fraction m of $J(n)$. We do not necessarily want to assume that $d_J + m = 1$ since it might take more than one time step for the juveniles to mature; in other words the difference $1-(d_J+m)$ will represent the fraction of juveniles at time $t = n$ that remain juveniles at time $t = n + 1$.
- Finally, the arrow marked "eggs" represents the number of juveniles produced by the adults laying eggs that hatch and produce new juveniles.[5] We will assume that each adult (female) lays some average number of eggs b per unit time that hatch and produce juveniles.

Note that the d_J, d_A, m, b parameters are all treated as *constants* here. That means in real-world terms that we are assuming whatever challenges or opportunities these organisms are encountering from their habitat are *not changing*. Models of this form should probably *only be used* when we can be reasonably sure that those assumptions are not being violated or when projections over short time spans are sufficient.

We record this more precise information in the updated diagram in Figure 8.2 Now we are ready to turn our understanding of the flows into and out of the compartments into a model. The idea is that we total up *all of the changes, both increases and decreases for each compartment*. Referring to Figure 8.2, we get the following

$$\begin{aligned} J(n+1) - J(n) &= b \cdot A(n) - m \cdot J(n) - d_J \cdot J(n) \\ A(n+1) - A(n) &= m \cdot J(n) - d_A \cdot A(n). \end{aligned} \tag{8.1}$$

[4]For instance, if the adults have better defenses against predators than the juveniles do.

[5]We are also assuming, in effect, that the time step is the time it takes for the eggs to hatch.

This can be simplified and rearranged to the final form in (8.2).

$$\begin{aligned} J(n+1) &= (1 - m - d_J) \cdot J(n) + b \cdot A(n) \\ A(n+1) &= m \cdot J(n) + (1 - d_A) \cdot A(n). \end{aligned} \tag{8.2}$$

This is what we meant before by a system of simultaneous difference equations. Each is linear and homogeneous and has constant coefficients according to the terminology introduced in Chapter 7. Note that the values $J(n+1)$ and $A(n+1)$ both depend on both $J(n)$ and $A(n)$ in this case.

As a final comment we mention that the constants b, m, d_J, d_A will usually depend on the species that we are looking at and also on the habitat. The constants d_J, d_A would include contributions representing predation from other species, the average life span of the individuals, etc. The constant b is determined by how successfully the species is reproducing; the constant m is related to how long the juvenile period lasts. Experimental data could be used to estimate these parameters and the process could be described as "fitting" a model of this form to the experimental data. This is somewhat analogous to what we did in Chapters 4, 5, and 6 to fit linear, exponential, and power law models. However, more sophisticated techniques are required for this sort of fitting. We will not discuss that aspect of developing such models in detail. △

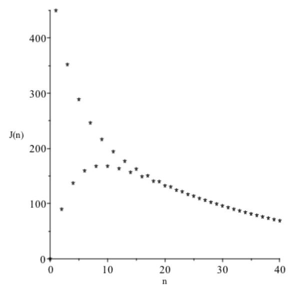
Generalizing what we did in Example 8.1, we can incorporate any number of age strata.

Example 8.2. A general population model for a population structured by age could be constructed in a similar fashion. Say we wanted to model a human female population broken down into 7 age groups:

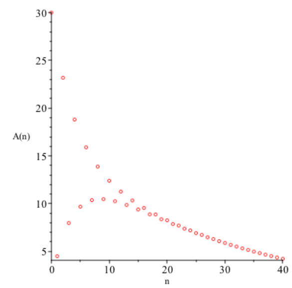
$$\begin{aligned} P_1(n) &= \text{number between 0 and 9 years in age at time } n \\ P_2(n) &= \text{number between 10 and 19 years in age at time } n \\ P_3(n) &= \text{number between 20 and 29 years in age at time } n \\ P_4(n) &= \text{number between 30 and 39 years in age at time } n \\ P_5(n) &= \text{number between 40 and 49 years in age at time } n \\ P_6(n) &= \text{number between 50 and 59 years in age at time } n \\ P_7(n) &= \text{number between 60 years and older at time } n \end{aligned}$$

As in the previous example, let us assume that all birth and death rates are constant. For simplicity, say the time step is 10 years, so if a woman counted in $P_5(n)$ survives for that 10 years, she will be counted in $P_6(n+1)$. Imagine a diagram as in Figure 8.2, but now with 7 boxes, one for each age group, arranged in order from left to right. For each box, we will have an arrow to the next one to the right corresponding to indivduals who survive and join the next age group. Next, each box has an arrow representing deaths. Say some fraction $d_i \cdot P_i(n)$ of the women in group i *do not survive* to enter the $(i+1)$st group, so $(1 - d_i) \cdot P_i(n)$ do contribute to $P_{i+1}(n+1)$. Finally, suppose that b_i is the birth rate per unit time (10 years) for the women in group i. This means that we also have arrows back from the box for P_i to P_1, given by $b_i P_i(n)$. We expect b_1 and b_7 to be very small, but probably not exactly zero.[6] Finally, in the last box for the 7th group, $(1 - d_7) \cdot P_7(n)$ represents

[6]With the decreasing average age of onset of puberty for girls, births to mothers of age 9 and younger have been recorded(!) Similarly, a few women remain fertile past the usual average age



(A) The juveniles.



(B) The adults.

FIGURE 8.3. The evolution of the juveniles and adults with $d_A = .85$ in Example 8.3.

the individuals who survive from the 60 and older group. We can think of that flow as an arrow from the box for $P_7(n)$ to itself. Putting this all together we have

$$
\begin{aligned}
P_1(n+1) &= b_1 \cdot P_1(n) + b_2 \cdot P_2(n) + b_3 \cdot P_3(n) + b_4 \cdot P_4(n) + b_5 \cdot P_5(n) \\
&\qquad + b_6 \cdot P_6(n) + b_7 \cdot P_7(n) \\
P_2(n+1) &= (1-d_1) \cdot P_1(n) \\
P_3(n+1) &= (1-d_2) \cdot P_2(n) \\
P_4(n+1) &= (1-d_3) \cdot P_3(n) \\
P_5(n+1) &= (1-d_4) \cdot P_4(n) \\
P_6(n+1) &= (1-d_5) \cdot P_5(n) \\
P_7(n+1) &= (1-d_6) \cdot P_6(n) + (1-d_7) \cdot P_7(n).
\end{aligned}
\tag{8.3}
$$

Note that this is slightly different from the model in Example 8.1 because, except for the final group, no one in group i at time $t = n$ remains in group i at $t = n+1$. We have, for instance

$$P_2(n+1) - P_2(n) = -P_2(n) + (1-d_1) \cdot P_1(n),$$

because every woman who was 10 to 19 years old at time $t = n$ has aged out of that group by $t = n+1$. The $-P_2(n)$ terms cancel, leaving the right-hand side from the second equation in (8.3). $\triangle$

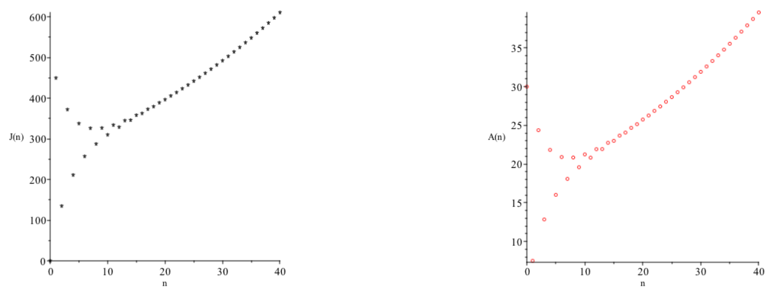
Now that we have looked at techniques for formulating structured population models as systems of difference equations, let's return to our first example and understand some of the behavior of solutions.

EXAMPLE 8.3. (a) To illustrate some possible behavior, let's consider the following (hypothetical) values of the constants in (8.2):

$$b = 15, m = 0.05, d_J = .9, d_A = 0.85$$

---

of onset of menopause and techniques like in vitro fertilization can be used to start pregnancies which then are brought to term after implantation of the fetus in an older woman. Births to mothers age 60 and above are possible, if extremely rare.



(A) The juveniles. (B) The adults.

FIGURE 8.4. The evolution of the juveniles and adults with $d_A = 0.75$ in Example 8.3.

As with a single difference equation, we need initial values to get the process started. So let's assume $J(0) = 0$ and $A(0) = 30$.[7] Starting from the initial conditions we compute:

$$J(1) = (1 - 0.05 - 0.9) \cdot 0 + 15 \cdot 30 = 450$$
$$A(1) = 0.05 \cdot 0 + (1 - 0.85) \cdot 30 = 4.5.$$

Then with the values of $J(1)$ and $A(n)$ in (8.2) we find

$$J(2) = (1 - 0.05 - 0.9) \cdot 4.5 + 15 \cdot 4.5+ = 90$$
$$A(2) = 0.05 \cdot 450 + (1 - 0.85) \cdot 4.5+ = 23.175,$$

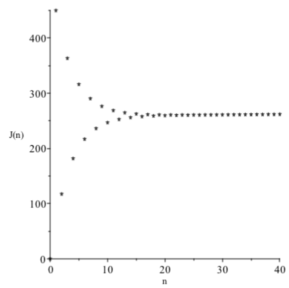
and so forth. We can compute as many different steps as desired. Let's not worry about the fact that we have fractional parts in some of the numbers here. We can compute the values out to $n = 40$ by using Excel in a way similar to what we did in Chapter 7 to compute approximate solutions of single difference equations. This can be done easily by setting up two parallel columns, say `A` for the J values, `B` for the A values. We enter the initial conditions in row 1 in those columns. Then in `A2` enter the formula for computing $J(1)$ in terms of $J(0)$ and $A(0)$: `= 0.05*A1 + 15*B1` and `= 0.05*A1 + 0.15*B1`. Then copy and paste those formulas down the rest of those columns. Since our addresses are not relative, the row numbers are increased by one in each subsequent row. You will see that there is an initial period of large oscillations from one time step to the next. But that appears to die down, and the populations of juveniles and adults tend to lower and lower values. Figures 8.3a and 8.3b show the evolution over time in this case.[8] Note that both portions of the population seem to be decaying to zero.

(b) On the other hand, suppose we change d_A to 0.75 (meaning that more adults survive to the next time step and hence more eggs are produced). Figures 8.4a and 8.4b show the evolution over time in this case. We see what looks like exponential growth (after an initial oscillatory period).

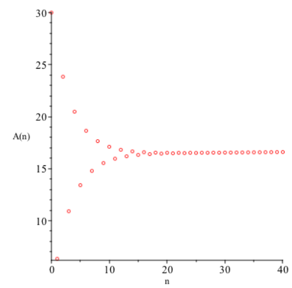
(c) Finally, with $d_A = 0.7895$, the numbers of juveniles and adults appear to be settling down to equilibrium values (again, after an initial oscillatory period). Figures 8.5a and 8.5b show the evolution over time in this case. We conclude this

[7] For instance, this could mean that 30 mosquitoes enter a new habitat with no juveniles and start to breed there.

[8] I used software other than Excel to generate these, of course.



(A) The juveniles.



(B) The adults.

FIGURE 8.5. The evolution of the juveniles and adults with $d_A = 0.7895$ in Example 8.3.

example by examining what we can say about equilibrium solutions of the system with $d_A = 0.7895$. As we know from Chapter 7, equilibrium solutions should come by taking the difference equation and putting it in the rearranged form giving the change in Q:

$$Q(n+1) - Q(n) = \text{right-hand side}.$$

Then we setting the right-hand side equal to zero and solve. If we do that with (8.1), we get the simultaneous system:

$$\begin{aligned} 0 &= -0.95 \cdot J(n) + 15 \cdot A(n) \\ 0 &= 0.05 \cdot J(n) - 0.7895 \cdot A(n). \end{aligned} \tag{8.4}$$

defining the equilbrium solutions.[9] If we examine the computed solutions we see

$$J(100) \doteq 258.3 \quad \text{and} \quad A(100) \doteq 16.4$$

are close to the apparent equilibrium values from Figures 8.5a and 8.5b.

It is easy to check that the right-hand sides of (8.4) evaluate to numbers very close to zero for these A and J values.[10] $\triangle$

---

[9]These two equations are *nearly* proportional. In fact given the values for m, d_J, b that we were using, there is a unique $d_A = \frac{15}{19} \doteq 0.7895$ for which this is true exactly. In this situation, there are infinitely many different solutions of a system of equations like (8.4).

[10]Technical note: It is possible to derive exact formulas for, and a complete quantitative and qualitative description of, these solutions $J(n)$ and $A(n)$ using *linear algebra*. What is needed is information about the *eigenvalues* α_1, α_2 of the matrices $\begin{pmatrix} .05 & 15 \\ .05 & 1 - d_A \end{pmatrix}$ from the right-hand sides of (8.2). It turns out that $J(n)$ and $A(n)$ have the form

$$J(n) = c_1\alpha_1^n + c_2\alpha_2^n \quad \text{and} \quad A(n) = d_1\alpha_1^n + d_2\alpha_2^n,$$

where c_1, c_2, d_1, d_2 are constants depending on the initial conditions and on d_A. As long as $\begin{pmatrix} c_1 \\ d_1 \end{pmatrix}$ is an eigenvector for α_1 and $\begin{pmatrix} c_1 \\ d_1 \end{pmatrix}$ is an eigenvector for α_2, then these formulas for $J(n)$ and $A(n)$ define solutions of the system of difference equations. When $d_A = 0.75$, one eigenvalue $\alpha_1 > 1$ and the other $\alpha_2 < 0$ with $|\alpha_2| < 1$, so we get approximate exponential growth for large n. The oscillatory behavior observed above comes from the alternating signs of α_2^n. When $d_A = 0.85$, $0 < \alpha_1 < 1$ and $\alpha_2 < 0$ with $|\alpha_2| < 1$, so both $J(n)$ and $A(n)$ tend to 0 for large n and we get oscillations again. Finally when $d_A = \frac{15}{19}$, $\alpha_1 = 1$ and $\alpha_2 < 0$ with $|\alpha_2| < 1$, which shows that there is an equilibrium which is approached as the oscillations from α_2^n die away.

8.3. Epidemiological Models

In this section we will consider some models for the spread of diseases through a population of some species of organisms. What we will see is that the general compartment setup can also be used to develop models of this sort. Most of the models in this section will be *non-linear*, though, because in order to spread, infectious diseases require some contact between an already-infected individual (or something that has been in contact with an infected individual[11]) and an individual who is not yet infected. We will see how modeling those sorts of interactions between populations can lead to non-linear terms in difference equation models.

The SIR Model. We begin with a simple example called the *SIR model.*[12] This epidemic model applies primarily to viral infections where the time scale for the infection and resolution (either by recovery or death) is much shorter than the life times of the affected individuals and infection *confers lasting immunity* in those who survive it. Typical time steps might be days or weeks, not years. Examples include measles, mumps, or rubella ("German measles"). These are *generally* non-fatal (especially in children). The basic model also applies to infections where death occurs in some fraction of those infected, but does not distinguish between recovery and death as an outcome. It does not apply well to diseases like the common cold, which do not confer immunity.

Assumptions 8.4. The assumptions of the SIR model include the following.

- The habitat is closed (recall from Chapter 7 that this means there is no in- or out-migration). Moreover, there is no important geographic or age structure of the population. Everyone can come into contact with everyone else.[13]
- The name comes from the fact that the population is divided into three groups, the *susceptible population*, S, the *infected population*, I, and the *removed population*, R. Individuals are removed if they contract the disease and either recover or die from it. It is a feature of the model that the total population $P = S(n) + I(n) + R(n)$ does not depend on n, which means that any individuals who do die are still being counted in the value of P.
- Susceptible individuals are those who have not contracted the disease yet. The basic SIR model does not attempt to distinguish between the individuals in the S group on the basis of better- or worse-functioning immune systems or other resistance to the disease.
- Infected individuals are assumed to be contagious the whole time they are infected, but they remain in contact with susceptibles.[14]

---

[11]Spoiler Alert: like the doll in the film *War for the Planet of the Apes* from summer 2017.

[12]The original form of this, a differential equation model, was proposed by W.O. Kermack and A.G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics". Proceedings of the Royal Society A. 115 (1927): 700–721.

[13]This basic model was first used, in fact, to understand outbreaks of disease in environments such as boarding schools, so these assumptions are more reasonable in that type of setting. But treating a portion of the human population of the world in isolation is unrealistic in most situations. With the greater mobility provided by easy, cheap air travel to virtually all parts of the world, there are new challenges to controlling disease outbreaks!

[14]That is, the model does not include any segregation or quarantine of infected individuals.

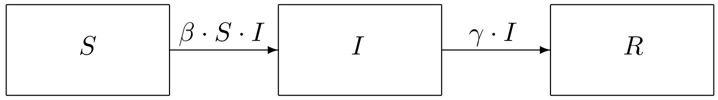


FIGURE 8.6. The SIR model in schematic form.

- New infections occur because of contacts between infected and susceptible individuals. The product $S(n) \cdot I(n)$ is a measure of the number of possible contacts. Some constant fraction of them produce new infections. That is, there is a constant β, $0 < \beta < 1$, such that the susceptibles are reduced by $\beta \cdot S(n) \cdot I(n)$ and the infecteds are increased by the same number.
- *Recovered* individuals acquire immunity and cannot be infected a second time. (The last part is also true for any individuals who die from the infection, of course!)
- The fraction of the infected group that is removed in each time unit is some constant γ with $0 < \gamma < 1$.

These assumptions are summarized in the compartment diagram from Figure 8.6. Translating this into difference equations we get

$$
\begin{aligned}
S(n+1) &= S(n) - \beta \cdot S(n) \cdot I(n) \\
I(n+1) &= I(n) + \beta \cdot S(n) \cdot I(n) - \gamma \cdot I(n) \\
R(n+1) &= R(n) + \gamma \cdot I(n)
\end{aligned}
\tag{8.5}
$$

EXAMPLE 8.5. Figure 8.7 shows the course of one epidemic modeled with SIR. To generate this, I used the hypothetical parameter values $\beta = 0.004$ and $\gamma = 0.11$. The initial conditions were $S(0) = 99$, $I(0) = 1$, $R(0) = 0$. We have a large susceptible group to start but the infection is introduced with $I(0) = 1$. With this combination of values for β, γ the susceptible population $S(n)$ steadily decreases to what appears to be an equilibrium value at a small positive number. The infected population $I(n)$ increases to a maximum at about $n = 21$, then decreases to values close to zero. Finally, the removed population $R(n)$ steadily increases until it levels off at something a bit than 100. The epidemic has died out but a small number of susceptible individuals are left. $\triangle$

The size of the removed population $R(n)$ is also effectively the cumulative number of infections in Figure 8.7. The cumulative number of infections in the 2014-2015 Ebola outbreak studied in Chapter 5 looks something like this.

Observations About SIR. Because of the non-linear $S(n) \cdot I(n)$ terms in two of the equations in (8.5), we should not expect to find exact formulas for the solutions of SIR models. However, we can make several qualitative observations based on the form of the equations.

- The constants β and γ have definite interpretations so their values could be estimated by observing the properties of an infectious disease even without some of the more advanced model fitting techniques to which we alluded above.
- The constant γ measures the *rate at which individuals move from the infected group to the removed group.* Hence $\frac{1}{\gamma}$ is a measure of *how long an*

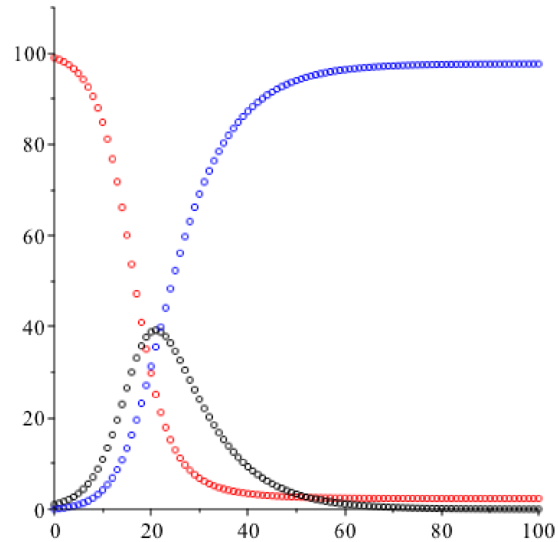


FIGURE 8.7. A computed solution of SIR Model–$S(n)$ in red, $I(n)$ in black, $R(n)$ in blue.

individual stays infected – the total time from infection to entry into the removed group.

- The constant β measures *how infectious the disease is* – how easily it spreads from contact. A high value of β would indicate a very infectious disease that would spread quickly; a small value of β would indicate a disease that is harder to catch in the sense that a smaller fraction of contacts produce infections.
- The second equation in (8.5) can be rearranged to the form:

$$I(n+1) - I(n) = (\beta \cdot S(n) - \gamma) \cdot I(n)$$

This says that $I(n+1) > I(n)$ exactly when $\beta \cdot S(n) - \gamma > 0$, or equivalently

$$\boxed{I(n+1) > I(n) \quad \Longleftrightarrow \quad S(n) > \frac{\gamma}{\beta}.} \tag{8.6}$$

This has several important consequences. If $S(0) > \frac{\gamma}{\beta}$ then the number of infected individuals will increase at first and this is the formal way an *epidemic* is defined. If $S(0) \leq \frac{\gamma}{\beta}$, on the other hand, then the number of infected individuals is decreasing from the start and no epidemic will occur. For this reason, epidemiologists refer to the constant

$$R_0 = \frac{\beta \cdot S(0)}{\gamma}$$

as the *basic reproduction number* for an infectious disease outbreak. Values $R_0 > 1$ indicate there will be an epidemic because $R_0 > 1$ implies $S(0) > \frac{\gamma}{\beta}$ as in (8.6). Values $R_0 \leq 1$ mean the outbreak will die out without producing an epidemic. Similarly, since $S(n)$ is steadily decreasing, we expect that as long as new infections occur, it will eventually be true that $S(n) = \frac{\gamma}{\beta}$. At this point, the change from $I(n)$ to $I(n+1)$ becomes *negative* and the outbreak begins to subside.

EXAMPLE 8.6. In the epidemic modeled in Example 8.5, in Figure 8.7 we can see that $I(n)$ begins to decrease at about $n = 21$. We used $\beta = 0.004$ and $\gamma = 0.11$ there, so $\frac{\gamma}{\beta} = 27.5$. In Exercise 7, you will check that $S(20) \doteq 29.8$ and $S(21) \doteq 25.2$. Therefore, $n = 21$ is indeed the first time that $S(n)$ drops below the threshold value $\frac{\beta}{\gamma} = 27.5$ and that is where the "corner is turned" and this epidemic begins to subside. $\triangle$

Extensions and Modifications. The basic SIR model can be modified to include other features and take other factors into account. This yields a large and growing class of different epidemiological models that are used to study the progress of disease outbreaks and manage public health responses.

- One very basic modification would be to treat the recovered individuals separately from the individuals who die as a result of a potentially fatal infectious disease. You will develop a modified SIR model of that form in Exercise 8.
- The SIR model could also be modified to account for population dynamics from births and deaths from other causes. This just amounts to adding new arrows in and out of the compartments in Figure 8.6.
- To remedy some of the unrealistic features of SIR from Assumptions 8.4, we know some diseases strike different age groups within a population in different ways because of differences in immunity, differences in overall health, etc. The basic SIR model can be combined with the techniques for age-structured populations that we developed in §2 of this chapter to study epidemiological questions within structured populations. You will see a simple example in Exercise 9.
- The SIR model can also be used to study the effects of vaccination on the dynamics of disease outbreaks. Recall that these models apply mainly to diseases where, in many cases, humans have learned to trick our immune systems into mounting the same defense responses acquired through actual exposure to the disease without undergoing the full exposure. You will see some examples in Exercise 10.
- More realistic models of the interaction between susceptibles and infecteds than the simple product term $S(n) \cdot I(n)$ have also be studied.

Epidemiological modeling is a huge subject and we have only scratched the surface. However, this limited exposure should convince you that the techniques we have developed have applications of great interest in public health questions.

8.4. Predator-Prey Models

The difference equation models that we have been studying in this chapter are very versatile and powerful tools for studying all sorts of dynamical problems. In this final section, we want to illustrate some of their power by modeling interactions between two different species. A well-known real world data set is the 90-year record of pelts of snowshoe hare and lynx bought by the Hudson's Bay Company

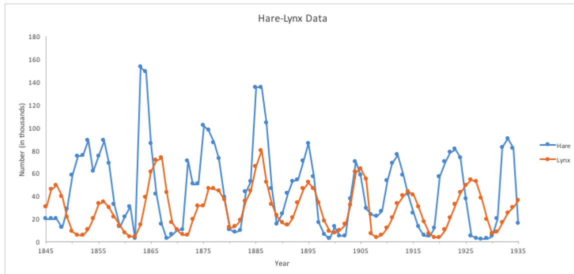


FIGURE 8.8. Hare and Lynx Pelts Bought by the Hudson's Bay Company.

in Canada. Human hunters caught animals of both species and sold their fur to the company, which recorded these numbers. But in the wild, the lynx (large cats) are predators who rely on the hares (large rabbits) for much of the food they consume and the hares are herbivores. The data is shown in Figure 8.8.[15] Now the idea is that the numbers here should reflect something about the *total populations* of the hares and the lynxes since the human hunting efforts were roughly comparable over the whole time. Note that there seem to *approximately cyclical variations* (together with a lot of "noise" that can probably be explained by variations in weather, random variations in the hare and lynx catches, and so forth) going on in this data set.

Our goal is to model the interaction between predators and prey. In rough terms, we can see that the presence of large numbers of prey animals should stimulate the growth of the population of predators (probably with some time delay in the response). Then in the next phase, large numbers of predators should drive down the prey population to the point where the large numbers of predators cannot be supported any longer. At this point the predator population should hit a maximum and then decrease. This will allow the prey population to recover and start another cycle. Any model we develop should ideally show something like this sort of cyclical variation, at least in some circumstances. We will consider a pair consisting of a *predator* species and the *prey* species that they feed on. Models of the same form can also be used for the interactions between parasites and hosts, between herbivores and the plant species they feed on, etc.[16]

ASSUMPTIONS 8.7. To formulate the basic form of our model, we make the following assumptions. We will denote by $P(n)$ the population of predators at time $t = n$ and by $Q(n)$ the population of prey individuals at time n.

- The predator species is entirely dependent on the prey species for food. The prey species, on the other hand, is entirely supported by other resources available in the habitat.
- The predator species population would follow an *exponential decay model* in the absence of any of the prey.

[15] From `http://jan.ucc.nau.edu/lrm22/lessons/predator_prey/predator_prey.html`, accessed July 20, 2017.

[16] The model we study here goes back to differential equation models introduced by Alfred Lotka and Vito Volterra in the 1920's.

- The habitat has ample resources for the use of the prey and the population of the prey species would follow a *exponential growth model* as in Chapter 7 in the absence of any of the predators.[17]
- Contacts between predators and prey are modeled, as in the SIR model, by product terms $P(n) \cdot Q(n)$.

These assumptions correspond to the following pair of first order, nonlinear difference equations where $0 < a < 1$ and $b, c, d > 0$ are constants:

$$\boxed{\begin{aligned} P(n+1) &= (1-a) \cdot P(n) + b \cdot P(n) \cdot Q(n) \\ Q(n+1) &= (1+c) \cdot Q(n) - d \cdot P(n) \cdot Q(n) \end{aligned}} \tag{8.7}$$

We will call this system of difference equations the *basic predator-prey model.*

Note that if $Q(n) = 0$ for all n, then there are none of the prey. In that case $P(n+1) = (1-a) \cdot P(n)$ is an exponential decay model and the predator population will go to zero as n increases. The closer a is to 1, the faster that exponential decay will be. On the other hand, if $P(n) = 0$ for all n, then $Q(n+1) = (1+c) \cdot Q(n)$. Since we assume $c > 0$, this gives exponential growth of the prey population and the larger c the faster the growth will be. The constant b in effect measures the value of the prey to the predators. Other things being equal, the larger b is, the more the prey will contribute more to the growth of the predator population. The constant d, on the other hand, measures how effective the predators are as predators. Other things being equal, the larger d is, the more the predators will contribute to the reducing the growth of the prey population.

Example 8.8. We consider the predator-prey model from (8.7) using the parameters $a = 0.4$, $b = 0.004$, $c = 0.02$, $d = 0.0006$ and initial conditions $P(0) = 4$, $Q(0) = 100$.[18] The solutions are plotted together in Figure 8.9. We can note several features of the solutions. First the overall trend seems to be a *cyclic variation* superimposed on an increasing trend in the heights of the peaks of both populations. This is perhaps somewhat surprising the first time you encounter solutions of the systems of this form. However, it does capture the some aspect of the cyclic variations that are observed in predator-prey interactions in the real world as in Figure 8.8(!)

A qualitative description of why the solutions of (8.7) exhibit this behavior is rather easy to provide. When the number of predators is low, the growth of the prey is nearly exponential. However, as the prey population grows, so does the $b \cdot P(n) \cdot Q(n)$ term in the first equation in (8.7). This produces a rise in the population of the predators in turn as we see in the "spikes" of the predator near $n = 40, 125$, and 225 in Figure 8.9. Around those spikes the $-d \cdot P(n) \cdot Q(n)$ terms are "outweighing" the exponential growth term in the second equation in (8.7) and the prey population is decreasing. But then the predators also decline to near zero and a new phase of growth in the prey population begins. $\triangle$

[17] As we know from Chapter 5, this is not realistic in the long run. However, these models can also be modified easily enough to make the prey species follow a logistic model if we wish. We will consider that sort of modification shortly.

[18] Technical note: For those familiar with the solutions of the Lotka-Volterra differential equation model, the sensitivity of the difference equation version to changes in these parameters and the variety of different sorts of solutions one can obtain can be quite surprising!

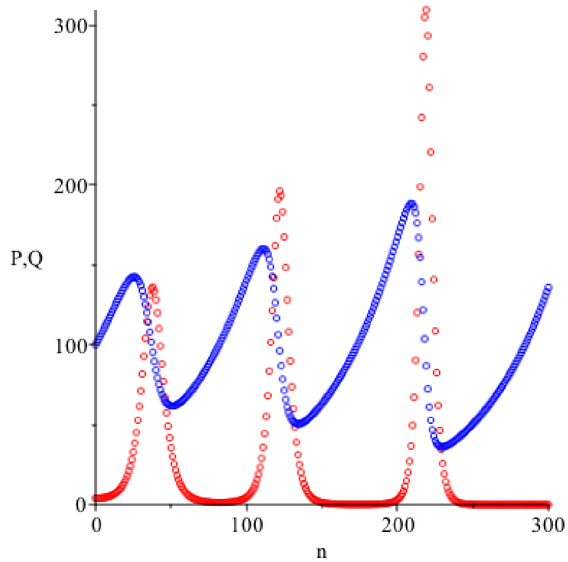


FIGURE 8.9. A computed solution of the predator-prey model–predators $P(n)$ in red, prey $Q(n)$ in blue.

Even when the behavior of one of these solutions is relatively well-behaved over the short run (as in this example), it is possible for $Q(n)$ to attain negative values at the low end of one of the oscillations, and then the solution can "go haywire" in spectacular ways. This happens in fact for $n = 875$ in our example and the results stop being realistic, or even realistically computable, shortly thereafter(!)) To avoid behavior of this type, and also (and probably more importantly!) to include more realistic assumptions in these models, it is also possible to build a *logistic growth model* for the prey into the set-up by considering *modified predator-prey models* of the form:

$$\boxed{\begin{aligned} P(n+1) &= (1-a)\cdot P(n) + b\cdot P(n)\cdot Q(n) \\ Q(n+1) &= (1+c)\cdot Q(n) - \frac{c\cdot Q(n)^2}{M} - d\cdot P(n)\cdot Q(n) \end{aligned}} \tag{8.8}$$

When $P(n) = 0$ for all n, the second equation becomes equivalent to (7.12), with M representing the carrying capacity for the prey.

EXAMPLE 8.9. We use the same parameter values $a = 0.4$, $b = 0.004$, $c = 0.02$, $d = 0.0006$ and initial conditions $P(0) = 4$, $Q(0) = 100$ as in Example 8.8, but introduce the new term $-\frac{0.02\cdot Q(n)^2}{300}$ with $M = 300$ in the second equation as in (8.8). The computed solutions are plotted in Figure 8.10.

As before, when the number of predators is low, the prey population is growing (but approximately according to a logistic model rather than exponentially). However, as the prey population grows, so does the $b\cdot P(n)\cdot Q(n)$ term in the first equation in (8.8). This produces a rise in the population of the predators in turn as we see in the "spikes" of the predator near $n = 50, 140$, and 230 in Figure 8.10. Around those spikes the $-d\cdot P(n)\cdot Q(n)$ terms are "outweighing" the growth term

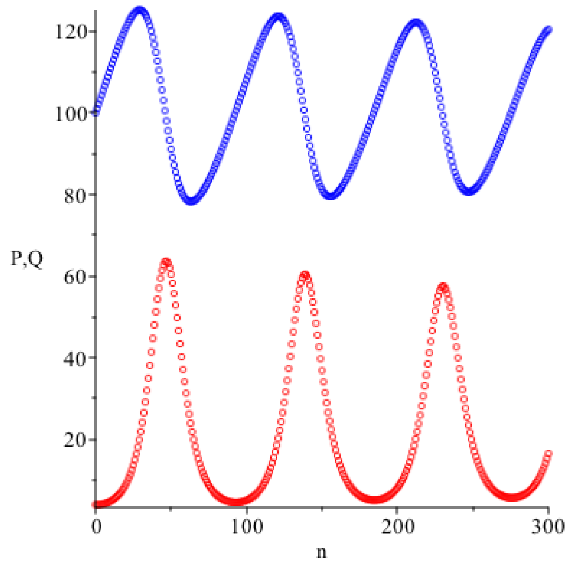


FIGURE 8.10. A computed solution of the modified predator-prey model (8.8)–predators $P(n)$ in red, prey $Q(n)$ in blue.

in the second equation in (8.8) and the prey population is decreasing. But then the predators also decline to near zero and a new phase of growth in the prey population begins. Here the oscillations in the predator and prey populations are "tamped down" and they even appear to be decreasing in amplitude as the cycles continue.

A different way to visualize what is happening here is to plot the data points $(P(n), Q(n))$ for $n = 0, \ldots, 300$. The result is called the *phase portrait* of the system of difference equations and is shown in Figure 8.11. In the phase portrait, the explicit dependence on the time $t = n$ is not shown, but we can see that as n increases the point $(P(n), Q(n))$ is spiraling clockwise and moving inward in the first quadrant of the P, Q-plane. $\triangle$

Considering what could be happening in Figure 8.11, it is natural to ask: Will the phase portrait continue to spiral, and if so, what is it spiraling toward? It is at least somewhat intuitively clear that if the spiraling continues for all n, then the phase portrait form of the solution is either tending toward some oval-shaped closed curve inside the region we plotted in Figure 8.11, or else toward a point (P_0, Q_0) defined by an equilibrium solution of the system (8.8). In fact, we can see that there is an equilibrium point in the interior of the first quadrant for all systems of the form either (8.7) or (8.8). To find equilibria of (8.7), we would solve

$$
\begin{aligned}
0 &= -a \cdot P(n) + b \cdot P(n) \cdot Q(n) = P(n) \cdot (-a + b \cdot Q(n)) \\
0 &= c \cdot Q(n) - d \cdot P(n) \cdot Q(n) = Q(n) \cdot (c - d \cdot P(n)).
\end{aligned}
$$

The first equation is zero if either $P(n) = 0$ or $Q(n) = \frac{a}{b}$. The second is zero if either $Q(n) = 0$ or $P(n) = \frac{c}{d}$. This means there are two equilibrium points: $(0, 0)$

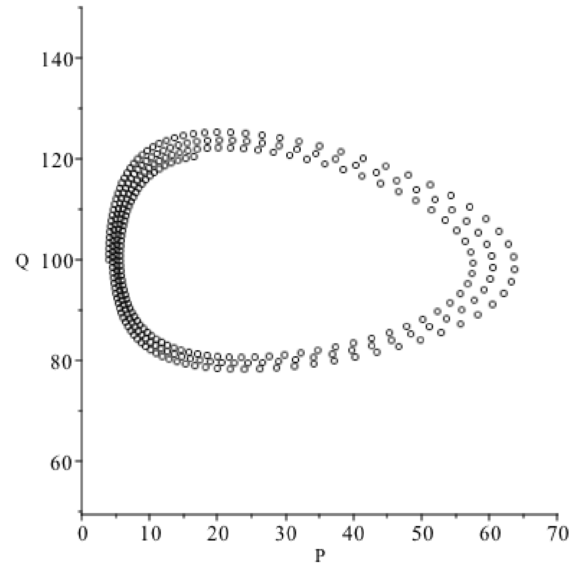


FIGURE 8.11. Phase portrait of the modified predator-prey model, solution from Figure 8.10.

where both predator and prey populations are zero, and the nonzero equlbrium

$$\boxed{\text{Equilibrium of (8.7) at } (P_0, Q_0) = \left(\frac{c}{d}, \frac{a}{b}\right)}$$

in the first quadrant. You will carry out the corresponding computation for the modified system (8.8) in Exercise 11. The result is that there are now three equilibria: two at the equilibria of the logistic model for the prey in the absence of the predators: $(0, 0), (0, M)$, and

$$\boxed{\text{Equilibrium of (8.8) at } (P_0, Q_0) = \left(\left(1 - \frac{a}{b \cdot M}\right) \cdot \frac{c}{d}, \frac{a}{b}\right).} \tag{8.9}$$

EXAMPLE 8.10. In Example 8.9, using the values of the parameters given there: $a = 0.4$, $b = 0.004$, $c = 0.02$, $d = 0.0006$ and $M = 300$, according to (8.9), the equilibrium point in the interior of the first quadrant is located at

$$\left(\left(1 - \frac{0.4}{0.004 \cdot 300}\right) \cdot \frac{0.02}{.0006}, \frac{0.4}{0.004}\right) \doteq (22.2, 100).$$

This is located inside the beginning of the spiral in Figure 8.11. If we plot more points, we see that it is apparently the case the solution is tending toward that equilibrium point as n increases without bound.[19] $\triangle$

[19]Technical note: This is rather awkward to explore using software like Excel since we need to go out to about $n = 1000$ to really start to see what is happening; the spirals are moving in quite slowly. Moreover that doesn't in itself show that the solution really does approach the equilibrium point, at least not to the level of evidence required by (pure) mathematicians(!) There is a theory of *linearization* around equilibria of nonlinear systems and asymptotic stability criteria that imply this really is what is happening in this case. The eigenvalues of the linearized system

Other sorts of interactions between two or more species can be modeled using similar techniques. We will not pursue those models in this elementary course, though.

8.5. Chapter Project

Background. Our chapter project for this chapter involves modeling the (short-term) global *carbon cycle*–the movements of carbon among the various living and nonliving components of the terrestrial environment. This includes processes acting over both long and short time scales. For instance, there are large stores of carbon in the rocky subsurface layers of the Earth's crust (in the forms of coal, oil and natural gas) that were created over long geological time spans by chemical processes acting on the tissues of plants that lived millions of years ago. We will not consider the process by which such deposits are formed because they take place over such long time scales. We will consider processes that take place over time spans of years or decades, though. This is the *short-term* part of the cycle.

The main reason this is important is that this short-term cycling of carbon is necessary for the continuation of life on Earth. This is true because carbon makes up about 50% of the *dry weight* of most plant and animal tissues (that is, what is left after the water is taken away). Anyone who has taken high school biology has surely marveled at the way plants and animals have co-evolved to use carbon in complementary ways. Animals breathe in oxygen from the atmosphere and exhale carbon dioxide as a product of their metabolisms from consuming plant tissues as food; they also return carbon to the atmosphere in other ways. On the other hand, plants take in carbon dioxide and incorporate it into their tissues by means of photosynthesis, while releasing oxygen back into the atmosphere as a byproduct.

As we mentioned in the Chapter Project on the Mauna Loa atmospheric carbon dioxide measurements data set from Chapter 4, in addition to being the primary constituent of living tissue, carbon is also present in the atmosphere (mostly in the form of carbon dioxide). Carbon dioxide is also an important *greenhouse gas* and the current life forms present could not exist on Earth without its warming effect. However, for the past 200 years or so humans have been increasing the atmospheric carbon dioxide content by means of fossil fuel burning and other activities. How will that additional carbon dioxide affect the existing mechanisms of the carbon cycle, Earth's climate and in turn its life forms? Exactly how the carbon cycle interacts with climate is a very difficult question and we will not attempt to model that because there are too many components and their interactions are only starting to be understood by scientists.

We will concentrate on understanding the carbon cycle itself and starting to understand how some human activities can create changes. A BIG disclaimer is certainly in order here: This is definitely a "toy" model that is much simpler than the real world and at the same time much simpler than the climate models that scientists are currently using to try to understand the evolution of the Earth's climate under the influence of anthropogenic sources of atmospheric carbon dioxide. For that reason, you should not take any of the computed values as especially realistic predictions.

---

at the equilibrium point are complex numbers with negative real parts, which implies that the solutions tend toward the equilibrium.

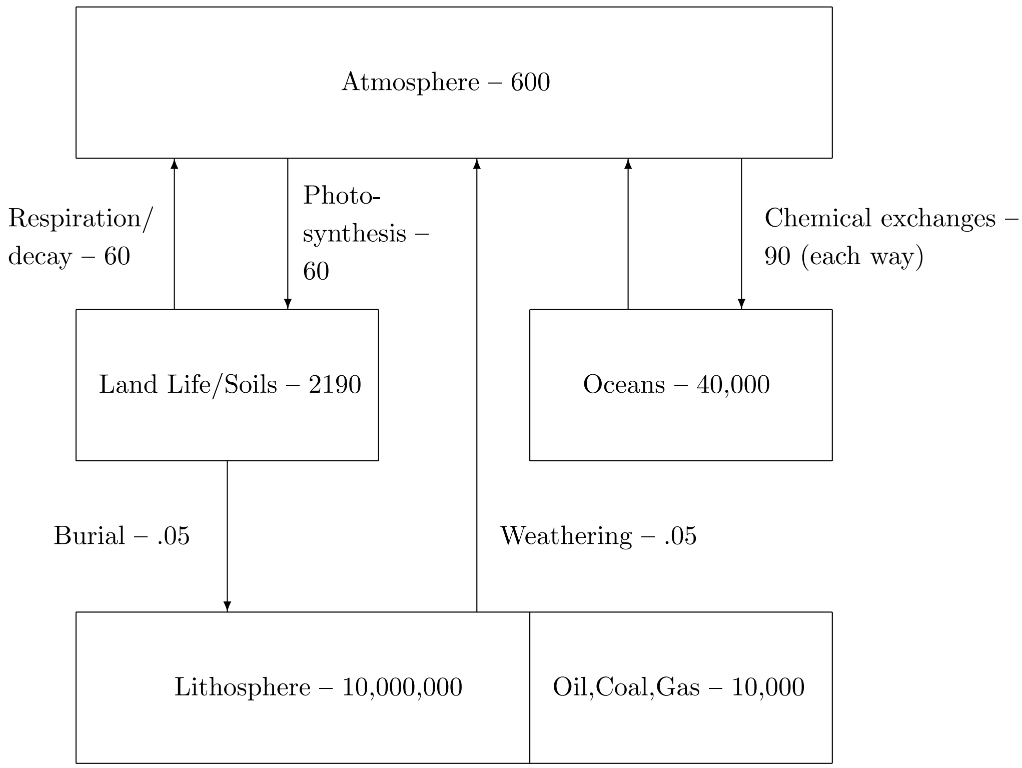


FIGURE 8.12. (Simplified) Carbon Cycle as of roughly 1750.

The Pre-Industrial-Revolution Carbon Cycle. Figure 8.12 shows an estimate of what the yearly carbon cycle looked like before about 1750 CE (this is an arbitrary date meant to predate the beginnings of large-scale industrial activity by humans making use of fossil fuel burning). The boxes show the major *carbon reservoirs* in the Earth system. The numbers in the boxes are total carbon contents in units of Gigatons = 10^9 metric tons = 10^{12} kilograms = 10^{15} grams.[20] The arrows show the *flows*, or "fluxes" of carbon between the different component systems, in units of 10^{12} kilograms *per year*.

This is a simplified schematic picture[21] because it does not take into account the fact that there are effectively two "layers" in the oceans, an upper layer where most aquatic life is located and where some carbon is dissolved into the water from the atmosphere, plus a lower layer containing much less life, but also large carbon-bearing sediments on the ocean bottoms. There are fluxes between those two layers as well that are thought to be possibly important for the dynamics of the carbon cycle over longer time scales. We are effectively "lumping those together."

[20] This is also called a "Petagram."

[21] Diagrams inspired by similar figures from the web site for a course at U. Arizona – `http://www.atmo.arizona.edu/students/courselinks/fall16/atmo336/lectures/sec3/carbon.html`, accessed July 21, 2017.

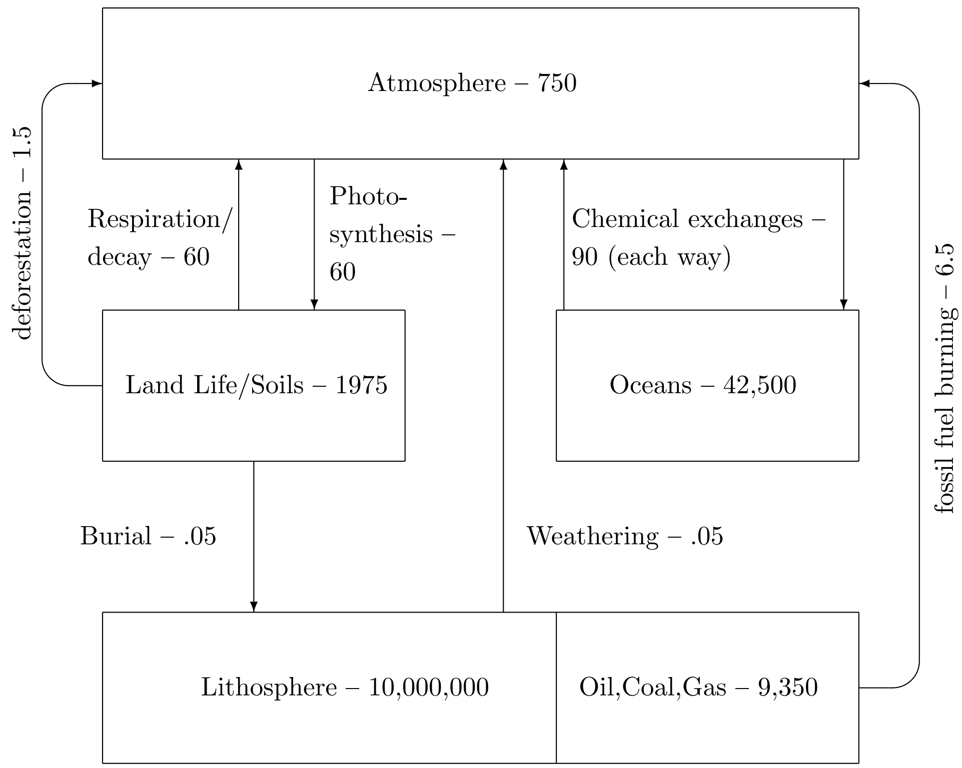


FIGURE 8.13. (Simplified) Carbon Cycle, as of roughly 1990.

Questions.

(A) Assume the fluxes in Figure 8.12 corresponding to Respiration/Decay, Photosynthesis, and the two-way Chemical Exchanges between the Atmosphere and the Oceans reservoirs, are all *proportional to the amounts in the source reservoir*, with proportionality constants determined by the values given. However the fluxes corresponding to Weathering and Burial are constant. Formulate a system of four difference equations equivalent to the diagram. It is thought that the carbon cycle was very nearly *at an equilibrium* before 1750. Is that true for the given data for your model?

(B) Now consider Figure 8.13 which shows estimates for the amounts of carbon in the various reservoirs and the fluxes *as of about 1990 CE*, so almost up to the present. Note this includes additional fluxes related to human activities that were not present before 1750 CE.[22]

Construct a new difference equation model for this regime, assuming that the fluxes (except for Burial and Weathering) that were present before are all proportional to the amount in the source reservoir as of 1990 with proportionality constants computed from the 1990 values of the fluxes.

[22] The numbers for pre-1750 and for 1990 are from different sources and were probably estimated using different methods. Hence they might not be completely consistent with each other.

However, the new, anthropogenic fluxes are determined like this: The fossil fuel burning and deforestation fluxes have been increasing by about 3% per year on average during this period. For instance, this means that the difference equation for the Atmosphere reservoir should contain terms like this corresponding to the flux in from the Oceans reservoir and the contribution from fossil fuel burning:

$$A(n+1) = A(n) + \frac{100}{41000} \cdot O(n) + \cdots + 6.5 \cdot (1.03)^n,$$

where n is the number of years after 1990. There will be other terms as well, indicated by the $\cdots$. Using the 1990 values of carbon content of the reservoirs as initial conditions, compute a solution for your model out to 2020 using Excel or suitable software.

(C) An atmospheric carbon amount of 600 Petagrams corresponds to an atmospheric carbon dioxide concentration of about 283 ppm (parts per million), and the atmospheric amounts of carbon are always proportional to the carbon dioxide concentration with the same ratio. From your computed solution in part (B), determine what your model says about the atmospheric carbon dioxide concentrations in ppm out to 2020. How does this compare with the Mauna Loa measurements we discussed in Chapter 4? Is the model prediction greater or smaller than the real-world data?

(D) (A "research/thought question.") Is the assumption that the fluxes are proportional to the amounts in the source reservoir somewhat reasonable (at least in the short-run)? One possible argument for arguing that the fluxes should at least be increasing as the amount of carbon in the Atmosphere reservoir increases would be this: More atmospheric carbon dioxide means more nutrients and better growing conditions for many plants, which means more food for animals to eat, which means more respiration and decay. Similarly, unless the Oceans are already saturated with carbon dioxide, the more there is in the Atmosphere, the more they can absorb and the more dissolved carbon dioxide there is, the more will be released. Does that seem to be consistent with the changes between Figure 8.12 and Figure 8.13? Are there other factors that might affect this, though?

(E) What is past is past and we cannot change it. But we can use models[23] to try to evaluate the effects of changes we might make in the future.
   (1) For instance, we might ask: Suppose we were able to limit fossil fuel burning to a constant level of 10 Gigatons per year and eliminate carbon emissions due to deforestation and other changes in land use entirely. (This is still significantly higher than the Kyoto Accords target of 5.7 Gigatons per year.) What would happen? To see, the easiest way (if you have not done this already) will be to create a new columns in your spreadsheet for the contributions due to those factors and use the values from them to compute the other quantities. You can manually change all values to 10 for the fossil fuel burning starting from any point.

[23]Preferably more realistic ones, of course.

(2) What if we continue tracking the model into the future keeping fossil fuel burning at 10 Gt per year? Does the CO2 concentration look like it will ever return to current levels? If so how long does it take?
(3) What is the largest constant (nonzero!) level for fossil fuel burning starting in 2017 that would still yield decreasing atmospheric CO2 levels by the year 2100? (This will require some experimentation!)
(4) Model validation "thought question" – Does your answer to (3) seem reasonable? Is there something going on in this model that might not be that realistic over these time scales? (Hint: Look at the Land Life/Soils figures. What would it take in real terms to have the amount of CO2 taken up by plants in photosynthesis increase by a factor of 10?)
(5) What are some of the potentially important features of the real world carbon cycle that are being left out of this model and that would need to be taken into account in order to produce a more realistic model? For example, might some of the fluxes in the cycle depend on other things like the global temperature, which we have not tried to address? (You may wish to look up information about this.)

Assignment. Write up the results of your investigations including the spreadsheet results and your answers to the questions above in a separate document.

Chapter Exercises

(1) In Example 8.3, we noticed that the model

$$J(n+1) = 0.05 \cdot J(n) + 15 \cdot A(n)$$
$$A(n+1) = 0.05 \cdot J(n) - 0.7895 \cdot A(n)$$

appeared to have a solution settling down to an equilibrium value when we used the initial conditions $J(0) = 0$, $A(0) = 30$. (See Figures 8.5a and 8.5b.)

(a) Does the solution tend to an equilibrium if we change the initial conditions? If so, do the equilibrium values of J, A depend on what the initial conditions are? If not, what happens?

(b) Now, keeping the initial conditions as before, investigate the effect of changing $b =$ the coefficient of $A(n)$ in the first equation in the system. What happens if $b > 15$? What happens if $b < 15$? Does that seem reasonable on biological grounds? Explain.

(2) In Example 8.2, we developed a model for the dynamics of a human population stratified into 7 groups by age.

(a) What would be reasonable values for the constants b_i? Recall these are birth rates per unit time for mothers in each of the age strata, but the time unit is 10 years. What would be reasonable values for the constants d_i? Recall these are death rates per unit time.

(b) Using Excel or other appropriate software, investigate the behavior of the solutions of the model (8.3) if

$$b_1 = 0.001, b_2 = 0.08, b_3 = 0.2, b_4 = 0.15, b_5 = 0.1, b_6 = 0.06, b_7 = 0.0001$$

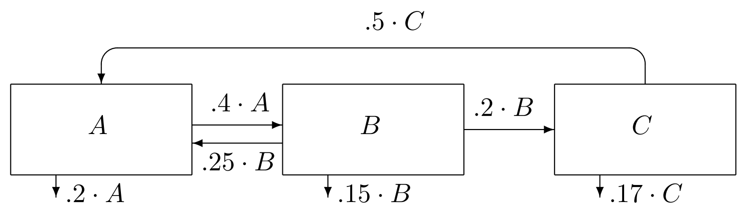


FIGURE 8.14. A three-compartment model.

and

$$d_1 = 0.003, d_2 = 0.004, d_3 = 0.006, d_4 = 0.008, d_5 = 0.05, d_6 = 0.1, d_7 = 0.25.$$

Track the solutions to $n = 20$ (at least) and describe the behavior qualitatively. In particular, is the total population undergoing growth, decline, or does it seem to be reaching an equilibrium?

(c) Repeat part (b), but changing the value of b_3 to 0.8. (Is that value unrealistically high? Think about the fact that the time step is 10 years(!))

(3) Convert the information from the compartment diagram in Figure 8.14 into a system of difference equations. For each compartment, assume that any portion of that compartment that does not flow out remains there in the next time step. For example, this means that the difference equation for A should have the form

$$A(n+1) = 0.4 \cdot A(n) + \cdots .$$

(4) This problem is adapted from information from the U.S. National Park Service.[24] Bulls (males) in the Denali caribou herd in Alaska can be divided into different age cohorts with different survival rates. Assume there are 1500 cows (females) in the herd in addition to the bulls[25] and each of them produces one calf each year. But only 50% of them are males and only 20% of them survive as long as one year because the very young male calves are vulnerable to wolves, eagles, and other predators. Young bulls between 1 and 4 years of age have grown to a size where they can protect themselves much more succesfully and have an annual survival rate about about 90%. They are not generally sexually active. Bulls between 5 and 7 years old are fully mature. For a period of 3 to 4 weeks each fall, during the annual "rut," they do not eat and spend much of their energy on an intensely competitive combat with other bulls. If they are successful, they may spend additional energy on mating with cows. As a result, they can lose one-third of their normal body mass each year and they spend the rest of the time after the rut season each year getting back to that normal mass to repeat the cycle. Because of this annual stress, a bull's annual survival rate is reduced to about 78% during this phase of their lives. After this, in years 8 and 9, survival declines steeply to about 15% per year. For the purposes of this problem, we will assume no bulls survive past age 10.

[24] From `https://www.nps.gov/articles/denali-caribou-herd.htm`, accessed July 19, 2017.

[25] Of course the real number of cows is also changing from year to year. But for the purposes of this problem we will make a simplifying assumption.

(a) Make a diagram with boxes representing the populations of bulls each age from 1 through 10 years. Put in arrows representing the additions and subtractions from each of those population groups each year.

(b) Convert your diagram from part (a) into a system of difference equations. Unlike most of the other models in this section, one of the difference equations in this system will *not be homogeneous*. Which one is it and why?

(c) Using Excel or other software, investigate the solutions of you system from part (b). If everything remained constant so that this model continued to hold, would the population of bull caribou be growing, declining, or tending to an equilbrium?

(5) (For readers of §3 of Chapter 7.) Show that a homogeneous linear 2nd order difference equation with constant coefficients, $Q(n+2)+r\cdot Q(n+1)+s\cdot Q(n) = 0$, is equivalent to the system of (first order) difference equations

$$\begin{aligned} P_1(n+1) &= P_2(n) \\ P_2(n+1) &= -r \cdot P_2(n) - s \cdot P_1(n), \end{aligned}$$

if we let $P_1(n) = Q(n)$ and $P_2(n) = Q(n+1)$.[26]

(6) Explain how the system of difference equations for the SIR model given in (8.5) is consistent with the assumption that the population is constant. (Hint: What is $S(n+1) + I(n+1) + R(n+1)$?)

(7) In this exercise you will verify the claims we made about the solution of the SIR model with $\beta = 0.004$ and $\gamma = 0.11$ studied in Examples 8.5 and 8.6.

(a) Using Excel or other software, generate the solution of the SIR model with these parameters and the initial conditions $S(0) = 99$, $I(0) = 1$, $R(0) = 0$. Generate plots like the ones from Figure 8.7 and compare.

(b) Examine the values $S(20)$ and $S(21)$ you computed in part (a). Explain why the number of infected individuals begins to decrease at $n = 21$.

(c) Repeat part (a) but changing the initial conditions to $S(0) = 20$, $I(0) = 1$, $R(0) = 0$. How is this solution different?

(8) In this exercise, you will develop a modified SIR model that tracks individuals who recover separately from those who die.

(a) The basic idea here is quite simple – we just want to split the R box in Figure 8.6 into two boxes, with one R_1 box representing the *recovered* individuals, and one R_2 box representing the individuals who *die* as a result of the infection. Develop an appropriate compartment diagram along these lines. In most cases the rate γ_1 at which individuals leave the I group and enter the R_1 group (recovered individuals) will be different from the rate γ_2 at which individuals leave the I group and enter the R_2 group of individuals who die. Your diagram should allow for that.

(b) Translate your diagram into a system of *four* difference equations for the functions $S(n), I(n), R_1(n), R_2(n)$.

[26]This idea can be developed to show that all second order and higher order difference equations can also be studied by means of the techniques introduced in this chapter.

(c) The first two equations in your model should be the same as those in (8.5), except the $-\gamma \cdot I(n)$ term in the second is replaced by something like $-(\gamma_1+\gamma_2)\cdot I(n)$. Take the values $\beta = 0.04$ and $\gamma_1 = 0.09$, $\gamma_2 = 0.02$, $S(0) = 99$, $I(0) = 1$, $R_1(0) = 0$ and $R_2(0) = 0$. Using Excel or other software, generate the solution of the modified SIR model with these parameters. Generate plots like the ones from Figure 8.7 and describe the course of the infectious outbreak.

(d) Refer to (8.6). What is the threshold value for epidemics with this model?

(9) In this exercise, you will develop a simple age-structured SIR model. Suppose for simplicity that we have a population P split into two strata by age (roughly, think "young" and "old" individuals). Each age stratum will contain susceptible, infected, and removed subgroups as in the basic SIR model. Call these $S_1, S_2, I_1, I_2, R_1, R_2$ The new feature is that infected people in either age stratum can infect others within either age stratum.[27]

(a) Develop an appropriate compartment diagram for this situation. There will be 6 boxes in all. To allow for differences in resistance to the disease, etc. allow for two different constants γ_1 and γ_2 governing the arrows from the infected boxes to the removed boxes. Also use four different constants β_{11}, β_{12}, β_{21}, and β_{22} in the arrows from the susceptible boxes to the infected boxes because contacts between people in different age strata might produce infections at different rates. For instance, you can think of β_{ij} as the constant for infections between susceptibles in stratum i and infecteds in stratum j.

(b) Translate your compartment diagram into a system of 6 difference equations. For example, the equation for S_1, the susceptibles in stratum 1 will look like this:

$$S_1(n+1) = S_1(n) - \beta_{11} \cdot S_1(n) \cdot I_1(n) - \beta_{12} \cdot S_1(n) \cdot I_2(n).$$

(c) Using Excel or other software, compute some solutions of these models. Use $\beta_{ij} = 0.004$ and $\gamma_i = 0.11$ for all $1 \le i, j \le 2$ and $S_1(0) = 99$, $S_2(0) = 100$, $I_1(0) = 1$, $I_2(0) = 0$, and $R_1(0) = R_2(0) = 0$.

(d) Change the parameters in part (c) letting $\beta_{21} = .01$ and $\beta_{22} = .007$ instead. What changes?

(10) With the basic SIR model, or modifications thereof, there are several ways we can study the effects of vaccination for an infectious viral disease. Imagine the situation of an isolated human population of size P.

(a) If we have the luxury of immunizing before any outbreaks occur, but we know the constants β and γ characteristic of the disease, how should public health authorities plan a vaccination campaign that will create enough "herd immunity" so that no epidemic outbreaks of the disease are possible? How many V out of the total population P need to be vaccinated? Not everyone needs to be vaccinated; how many vaccinations are necessary?

[27] If not, if for instance young people could only infect other young people, we would essentially have two separate SIR models.

(b) Once an outbreak is underway, it is still possible to attempt to slow it down or drive it to conclusion via a vaccination campaign. One approach would be to vaccinate some fixed number v_0 of the susceptible group each time step. How could the basic SIR model be modified to take this into account? Construct a new modified compartment diagram representing this situation and translate it into a system of difference equations.[28] (Hint: This is just adding one more arrow to the compartment diagram.)

(c) Another approach besides part (b) would be to vaccinate some fraction of the susceptible population each time step. How could the basic SIR model be modified to take this into account? Construct a new modified compartment diagram representing this situation and translate it into a system of difference equations.

(11) Show that the modified predator-prey system (8.8) has equilibria at $(0, 0)$, $(0, M)$ from the logistic model for the prey when no predators are present, *and* at

$$(P_0, Q_0) = \left(\left(1 - \frac{a}{b \cdot M}\right) \cdot \frac{c}{d}, \frac{a}{b}\right)$$

in the interior of the first quadrant.

(12) Refer to Example 8.9.

(a) Verify the calculations presented in that example using Excel or other suitable software.

(b) What happens to the computed solutions if the value of the coefficient a is changed to $a = .8$ (meaning that the predators die out faster if no prey are present)?

(c) What happens to the computed solutions if $a = .1$?

[28]Eventually the number of susceptible individuals will be reduced to negative values with this method. When that happens, the solutions of the model stop being realistic.

III. Data Analysis and Statistics

CHAPTER 9

Descriptive Statistics

9.1. Introduction

In the previous chapters, we have concentrated on modeling. In this section of the text, we will introduce a number of additional techniques for analyzing the information contained in datasets. We will use these even more extensively in later chapters in learning how to make inferences from data. Some of the basic questions involved in analyzing the distribution of values in a dataset are:

- Where is the *center, or middle* of the data?
- How *spread out* is the data?

Some important types of inferences we will learn about making are:

- Is a particular data value an *outlier*, and is it so unrepresentative that we might be justified in treating it separately or even omitting it from consideration?
- If we select a sample from some population of organisms and make some measurement on each individual (for example a body length or weight), can we estimate the average over the whole population, or perhaps a likely interval of values for the average of that measurement on the whole population?
- If two populations are compared by drawing samples and computing some statistic from measured values, is there a *demonstrable difference* between the populations (that is, a difference that is unlikely to come just from random variations stemming from the process of selecting the samples)?

In this chapter, we will present some first *descriptive statistics* that can be used to understand the distribution of a set of data values and that are the basis for the sorts of inference procedures we will address later. In general, by a *statistic*, we will mean a numerical value computed in some definite way from the data values. We have already seen some first steps in this kind of thinking in our discussion of least squares regression from Chapter 4, and also in our treatment of fitting power law distributions from Chapter 6.

Recall that in Chapter 4 we discussed how the straight line best fitting a bivariate dataset (x_i, y_i) for $i = 1, \ldots, N$ could be found by solving the *normal equations* from (4.6):

$$\left(\sum_{i=1}^{N} x_i^2\right) \cdot m + \left(\sum_{i=1}^{N} x_i\right) \cdot b = \sum_{i=1}^{N} x_i y_i$$

$$\left(\sum_{i=1}^{N} x_i\right) \cdot m + N \cdot b = \sum_{i=1}^{N} y_i.$$

If we divide through by N in the second equation, we obtain:

$$\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right)\cdot m + b = \frac{1}{N}\sum_{i=1}^{N} y_i.$$

You probably recognize the

$$\frac{1}{N}\sum_{i=1}^{N} x_i, \quad \text{and} \quad \frac{1}{N}\sum_{i=1}^{N} y_i$$

as the *numerical averages* of the x_i and the y_i. In this chapter we will introduce the notation

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

for the average, or *mean*, of the x_i and we will see how it is used as one sort of way to estimate the *center, or middle* of a distribution of data values. If we do this for the y_i as well, we get a mean $\overline{y}$ of the y_i, and then the second normal equation just says $m\cdot\overline{x}+b=\overline{y}$. In particular, the *point of averages* $(\overline{x},\overline{y})$ lies on the best-fit regression line $y=\hat{m}\cdot x+\hat{b}$ when we solve the normal equations to find the slope and intercept $\hat{m},\hat{b}$.

In earlier chapters, we also discussed (in a quite informal way) how the R^2 statistic could be used to evaluate the goodness of fit of a regression line. We did not discuss how the R^2 statistic was computed before, but we can now with the concepts introduced in the previous paragraph. In the case that we are modeling y as a linear function of an exactly known *explanatory* variable x, then the R^2 statistic is computed with this formula

$$R^2 = \frac{\left(\displaystyle\sum_{i=1}^{N}(x_i-\overline{x})\cdot(y_i-\overline{y})\right)^2}{\displaystyle\sum_{i=1}^{N}(x_i-\overline{x})^2\cdot\sum_{i=1}^{N}(y_i-\overline{y})^2}, \tag{9.1}$$

or by one of several possible rearranged forms. Taking square roots gives the so-called *correlation coefficient*, R.

In this chapter we will introduce a number of other techniques for analyzing data sets with different properties and build up to understanding the rationale behind the correlation coefficient (presented in a final, optional section).

9.2. Measures of the Center of a Data Set

We will begin with the question of how to estimate what we would think of intuitively as the middle, or center, of a collection of data values. There are two commonly-used statistics for this.

The first statistic we will discuss is the *mean*, or numerical average of the values x_i, $i=1,...,N$, defined as follows:

$$\boxed{\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{N}(x_1+\cdots+x_N).} \tag{9.2}$$

The mean is often used, for instance, to compute things like class averages on tests or other assignments.

Another way to say where the middle of a distribution of data values is by means of another statistic you have probably encountered, called the *median*. Roughly speaking, the median is a number such that half the data values are greater than or equal to that number and half the data values are less than or equal to that number. To simplify computing the median by hand, we sort the data values into non-decreasing order:

$$x_1 \le x_2 \le x_3 \le \cdots \le x_{N-1} \le x_N. \tag{9.3}$$

Then we can find the median just by looking for the "middle value."

If $N = 2K + 1$ is an *odd number* then there is a value right in the middle of the list, which is then the median because of the way the data values have been listed. This is the value x_{K+1} since there are K values $x_1, \dots, x_K$ less than or equal to x_{K+1} and also K values $x_{K+2}, \dots, x_{2K+1}$ values greater than or equal to x_{K+1}.

On the other hand, if $N = 2K$ is an *even number*, then there is no value exactly in the middle of the list, and the median is typically taken to be the average or mean of the two data values closest to the middle of the list: $\frac{x_K + x_{K+1}}{2}$.

Putting together these two cases, if the data is arranged in non-decreasing order as in (9.3), then

$$\boxed{\text{median} = \begin{cases} x_{K+1} & \text{if } N = 2K+1 \text{ is odd} \\ & \text{and} \\ \dfrac{x_K + x_{K+1}}{2} & \text{if } N = 2K \text{ is even.} \end{cases}} \tag{9.4}$$

(You can remember this as a formula if you think: take the middle of the list when N is odd and the average of the end of the first half of the list – the x_K – and the beginning of the second half of the list – the x_{K+1} – when N is even.)

Example 9.1. To keep things simple, let us consider two hypothetical data sets.

(a) First, consider these x_i:

$$5.2 \quad 8.9 \quad 2.1 \quad 7.5 \quad 9.3 \quad 4.3 \quad 4.7,$$

where $N = 7$, an odd number. The mean is given by

$$\overline{x} = \frac{5.2 + 8.9 + 2.1 + 7.5 + 9.3 + 4.3 + 4.7}{7} = \frac{42.0}{7} = 6.0.$$

If we rearrange to non-decreasing order as in (9.3), we have

$$2.1 \le 4.3 \le 4.7 \le 5.2 \le 7.5 \le 8.9 \le 9.3$$

and we see the middle number – the median – is 5.2.

(b) Second, consider these x_i:

$$5.2 \quad 8.9 \quad 2.1 \quad 7.5 \quad 9.3 \quad 4.3.$$

This is the same as the previous list, but omitting the 4.7, so now $N = 6$. We have

$$\overline{x} = \frac{5.2 + 8.9 + 2.1 + 7.5 + 9.3 + 4.3}{7} = \frac{37.3}{6} \doteq 6.2.$$

To find the median, we omit the 4.7 from the sorted list:

$$2.1 \le 4.3 \le 5.2 \le 7.5 \le 8.9 \le 9.3$$

and we average the number at the end of the first half and the number at the beginning of the second half:

$$\text{median} = \frac{5.2 + 7.5}{2} = 6.35.$$

We will discuss some patterns visible in these examples next. △

A first observation is that in both cases the mean and the median *are different*. This is to be expected because these two statistics give a measure of the middle of a collection of data values in two different ways.

A physical analogy may help to make what the *mean* does more vivid. Suppose we have a collection of *equal*, say 1 kg, masses located along a beam. We introduce a reference point and measure distances from that point in either direction using some distance unit, say meters. Then the location of each mass is described by some real number x_i and the reference point is located at the point corresponding to the number 0. Reversing the process, given any data set, we can also visualize the numbers involved as a collection of positions of a collection of equal masses in this way.

Then, by (a special case of) the *law of the lever* introduced by Archimedes,[1] the configurations will *balance exactly* if the beam is placed on a fulcrum at the point corresponding to the mean, $\overline{x}$. For instance, in Example 9.1 (a), the collection of equal masses would balance with a fulcrum at $\overline{x} = 6.0$, and that balance point shifts to $\overline{x} = 6.2$ when the mass at location $x_i = 4.7$ removed. In effect, the balance point, *the mean, shifts to the right* (i.e. to a larger numerical value) because we have taken away one of the masses lying to the left of (i.e. smaller than) the original balance point. The same kind of shifting of the mean to the right would occur if we moved any one of the numbers in (a) farther to the right. For instance, increasing the largest number from 9.3 to 10.3 shifts the mean to $\frac{43.0}{7} \doteq 6.15$. On the other hand, removing a value lying to the right of (i.e. larger than) the mean, or shifting one or more of the values to the left (i.e. to a smaller value) would make the mean smaller.

The description of the median leading up to the recipe for computing it in (9.4) shows that the median estimates the middle in a quite different way. In particular, notice that moving a data value to the left or right can often have no effect at all on the median of a data set. For instance, if we shifted the value 9.3 to 10.3 in the data set from part (a) of the example, the position of the median would not change at all; we would still have median = 5.2. Similarly, if we shifted any of the four numbers not involved in the computation of the median in part (b), but we left the ordering unchanged, there would be no effect on the median. In this sense, the median is more *robust* than the mean – it is less affected by small changes in the data values.

Putting these observations together, we can see that the mean takes the individual values in the data set more into account than the median does, while the median has more to do with the relative ordering of the data values. There are

[1] In his *Equilibrium of Planes*, Propositions 6 and 7. This is also the background for the much-discussed boasting quotation traditionally ascribed to Archimedes: "Give me (a lever and) a place to stand and I will move the Earth!"

situations where each of these ways of finding the middle is more appropriate, and we might even want to use both of them in combination(!)

Example 9.2. In part (a) of Example 9.1, we might say that the distribution is *skewed* towards larger values, meaning that the larger values are farther from the median (on the right side) than the smaller values are from the median (on the left side). We can see this, for instance by comparing the difference $9.3 - 5.2 = 4.1$ (from the largest value) with the difference $5.2 - 2.1 = 3.1$ (from the smallest value) and so on. And in fact the median is *smaller than the mean* because those larger values pull the mean in that direction.

On the other hand in (b) of the example, the median is *larger than the mean.* When that happens it is sometimes claimed that the data values should be *skewed* towards smaller values. This is partly true in this case if we compare the difference $9.3-6.35 = 2.95$ (from the largest value) with the difference $6.35-2.1 = 4.25$ (from the smallest value). However the pattern is not quite as consistent as in part (a). For instance, the second largest value 8.9 is actually farther from the median than the second smallest value 4.3 is. $\triangle$

We will discuss a way to quantify the notion of skewness of a distribution later in the chapter and we will return to these examples to illustrate that the connection between skewness and the relative positions of the median and mean is more complicated than is sometimes claimed(!)

9.3. Measures of Spread and Distribution

The Range. One of the most basic notions of how spread-out a data set is simply the numerical range of values represented in the data. This can be measured by computing the largest and smallest values, say *max* and *min*, and taking the difference

$$\boxed{\text{range} = max - min} \tag{9.5}$$

Example 9.3. The range of the datasets in both parts of Example 9.1 is

$$\text{range} = 9.3 - 2.1 = 7.2.$$

Note that the other five or six data values do not enter here at all. $\triangle$

The Standard Deviation. To capture the difference between a data set like the one from part (a) of Example 9.1

$$5.2 \quad 8.9 \quad 2.1 \quad 7.5 \quad 9.3 \quad 4.3 \quad 4.7$$

with $\overline{x} = 6.0$, and another like

$$6.2 \quad 7.9 \quad 2.1 \quad 6.5 \quad 9.3 \quad 4.8 \quad 5.2,$$

which also has $\overline{x} = 6.0$ and range $9.3-2.1 = 7.2$, we can use another statistic known as the *standard deviation*, or SD for short, usually denoted by s. By definition, the SD is the *root mean square deviation from the mean.* To see what this means (and see how to compute s by hand), we will break this down. This will also show why s is a reasonable measure of spread.

- The *deviation* of a data value x_i from the mean $\overline{x}$ is just $x_i - \overline{x}$.

- The *square deviation* is $(x_i - \overline{x})^2$. (The effect of squaring is to make larger deviations count more and smaller ones count less, while disregarding whether the deviation is positive $x_i > \overline{x}$) or negative $(x_i < \overline{x})$.)
- The *mean square deviation* is essentially the mean of the square deviations. And in fact some (mostly older) treatments of statistics would use exactly the mean of the square deviations, in the form
$$\frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2.$$
However, for technical reasons,[2] we will use a slightly different version with the N in the denominator replaced by $N - 1$.
$$\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2.$$
- Finally, the *root mean square deviation* means to take the square root of the result of the previous step.

Putting everything together, we get the following formula for the standard deviation:

$$\boxed{s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}.} \tag{9.6}$$

One reason for taking the square root is so that we obtain a quantity that is measured *in the same units* as the original data values x_i. Note, for instance, that if x_i were distances in units of meters (m), then the square deviations $(x_i - \overline{x})^2$ would have units of m^2. To compare our measure of spread with the values of the data, it is more convenient to have something in the same units, and because of the square root, s would have units of m again.

There is also a rearranged form of this equation that is somewhat more efficient for hand computation. We will be relying on software such as Excel for these computations, though, so we will address this only in the exercises.

EXAMPLE 9.4. Consider the hypothetical data sets

$$5.2 \quad 8.9 \quad 2.1 \quad 7.5 \quad 9.3 \quad 4.3 \quad 4.7$$

and

$$6.2 \quad 7.9 \quad 2.1 \quad 6.5 \quad 9.3 \quad 4.8 \quad 5.2$$

given above. Both have $\overline{x} = 6.0$ and both have the same range. We now compute the standard deviations. For the first one we compute the square deviations:

$$(5.2 - 6.0)^2 = 0.64, (8.9 - 6.0)^2 = 8.41, (2.1 - 6.0)^2 = 15.21, (7.5 - 6.0)^2 = 2.25$$
$$(9.3 - 6.0)^2 = 10.89, (4.3 - 6.0)^2 = 2.89, (4.7 - 6.0)^2 = 1.69.$$

Hence the mean square deviation (using a denominator $N - 1 = 6$) is

$$\frac{0.64 + 8.41 + 15.21 + 2.25 + 10.89 + 2.89 + 1.69}{6} \doteq 6.997$$

[2]Technical note: This is done to obtain an *unbiased estimator* for the population standard deviation when we sample from a population with a known distribution.

and hence

$$s \doteq \sqrt{6.20} \doteq 2.645.$$

Similarly, for the second data set we compute

$$(6.2-6.0)^2 = 0.04, (7.9-6.0)^2 = 3.61, (2.1-6.0)^2 = 15.21, (6.5-6.0)^2 = 0.25$$
$$(9.3-6.0)^2 = 10.89, (4.8-6.0)^2 = 1.44, (5.2-6.0)^2 = 0.64.$$

Hence the mean square deviation (using a denominator $N - 1 = 6$) is

$$\frac{0.04 + 3.61 + 15.21 + 0.25 + 10.89 + 1.44 + 0.64}{6} \doteq 5.347$$

and hence

$$s \doteq \sqrt{5.347} \doteq 2.312.$$

The values in the second data set (other than the maximum and the minimum, which are the same in both data sets) are somewhat more closely "clumped" together around the mean. This can be seen, for instance, by locating the points on a number line. The smaller value of s for the second data set is an indication of this feature of the two distributions. $\triangle$

Once we have the mean $\overline{x}$ and the standard deviation s, we can use those numbers to define a collection of standard intervals that are often used to describe the distribution of the data values. For instance, the interval defined by $\overline{x} - s \leq x \leq \overline{x} + s$ is called the interval of x values *within one standard deviation of the mean.* Similarly, the interval $\overline{x} - 2 \cdot s \leq x \leq \overline{x} + 2 \cdot s$ consists of x-values *within two standard deviations of the mean.*

EXAMPLE 9.5. For instance, with the first data set from Example 9.4, we have $\overline{x} = 6.0$ and $s \doteq 2.645$. So the interval of values within two standard deviations of the mean is the interval of x with

$$0.71 = 6.0 - 2 \cdot 2.645 \leq x \leq 6.0 + 2 \cdot 2.645 = 11.29.$$

Note that *all* of the data values are in this interval in this case. $\triangle$

Chebyshev's Theorem and One Criterion for Outliers. There is a theoretical result known as Chebyshev's theorem[3] that states that *in every data set, at least*

$$\left(1 - \frac{1}{n^2}\right) \cdot 100\%$$

of the data values will be in the interval within n standard deviations of the mean. With $n = 2$, for instance, this gives at least $\left(1 - \frac{1}{4}\right) \cdot 100\% = 75\%$ within two standard deviations of the mean. Thus it is perhaps not surprising that a large fraction of the data values ended up in this interval for our example data set. Similarly, with $n = 3$, the interval of values within *three* standard deviations of the mean will always contain at least $\left(1 - \frac{1}{9}\right) \cdot 100\% \doteq 89\%$ of the data values. The statement of Chebyshev's theorem also gives a sort of indication why the *standard*

[3]Named after the Russian mathematician Pafnuty Chebyshev, 1821-1894. Because this is a Russian name, it was originally written in the Cyrillic alphabet. It has been transliterated to the Roman alphabet to give an English equivalent. There are several different systems for these transliterations and you will find many different-looking forms of the name if you look at different sources: Chebysheff, Tchebyshev, etc.

deviation has that name: there are uniform statements one can make about the distribution of the values in all data sets using the values of s (and $\overline{x}$)!

These estimates from Chebyshev's theorem are, in a sense, "worst case" lower bounds. Under some fairly general conditions, it can be shown that this fraction of data values within two or three standard deviations is actually much larger – often more like 95% within 2 SDs and 99% within 3 SDs.[4] This is the basis for one rule of thumb that is often used to identify *outliers* in a data set. This criterion is to consider any value that lies more than three standard deviations from the mean (in either direction) to be an outlier. This criterion is sometimes called the "3σ rule" since the Greek letter σ ("sigma") is often used to represent a parameter of theoretical distributions analogous to the data set SD, which we have denoted by s. The question then becomes, what should be done with outliers? Simply removing them and proceeding as if those values were never included in the data is generally considered to be a "shady" practice because it is essentially hiding information that might in fact be valid and might lead to different conclusions. That sort of "cherry-picking" is not an honest use of data but it is done from time to time; statistics can definitely be *misused!* Labeling the questionable data values as outliers but still considering their effects in the analysis is considered better statistical practice in most cases.

The Quartiles. The basic idea here is to extend the way the median of a data set sorted in non-decreasing order is defined and identify two additional values:

- Q_1, the *first quartile*, at the middle of the first half of the data set (so, roughly speaking, one quarter of the data values lie at or below Q_1), and
- Q_3, the *third quartile*, at the middle of the second half of the data set (so, roughly speaking, three quarters of the data values lie at or below Q_3).

The median then serves as the *second quartile*, with one half of the data values at or below the median. Perhaps unfortunately, and certainly inconveniently, there is no standard way to handle the different special cases that come up depending on whether the first and second halves of a data set have even or odd numbers of entries. In this text we will use the following method to find the first, second, and third quartiles. The guiding principle is that the median is *included* in the two halves used to compute the first and third quartiles when it is one of the values in the data set.[5]

(1) First compute the median as before, giving the second quartile.
(2) If the original data set had an odd number $N = 2K + 1$ of elements, then include the element x_{K+1} in both halves. The first half will be $x_1, \ldots, x_{K+1}$ and the second half will be $x_{K+1}, \ldots, x_{2K+1}$, each with $K+1$ elements.
(3) If the original data set had an even number $N = 2K$ of elements, then make the first half $x_1, \ldots, x_K$ and the second half $x_{K+1}, \ldots, x_{2K}$.

[4]Technical note: This is true for instance if the data is *normally distributed*, as we will see later.

[5]We will show later how to find the quartiles following this convention using Excel. But you should be aware that other books and other software packages sometimes use different procedures to find the first and third quartiles. For example, sometimes when N is odd, the median *is not included in either half* to compute Q_1 and Q_3. To check numerical results, you will need to know what convention is being used.

(4) Then Q_1 is the median of the first half and Q_3 is the median of the second half, computed as we described before.

Here are examples illustrating some of the possible cases.

EXAMPLE 9.6. (a) First consider the sorted (hypothetical) data set

$$1.1 \le 2.3 \le 2.4 \le 4.5 \le 5.6 \le 7.8 \le 9.5 \le 10.3$$

Here $N = 8$, which is even. The median (second quartile) is $\frac{4.5+5.6}{2} = 5.05$. The first quartile is then computed as the median of the first half:

$$1.1 \le 2.3 \le 2.4 \le 4.5,$$

so $Q_1 = \frac{2.3+2.4}{2} = 2.35$, since the first half also has an even number of elements. Similarly, $Q_3 = \frac{7.8+9.5}{2} = 8.65$. Something similar would happen for data sets of length $N = 4, 12, 16$, etc. as well.

(b) Now consider another hypothetical sorted data set with $N = 7$ elements

$$3.8 \le 3.9 \le 4.3 \le 5.4 \le 6.7 \le 8.9 \le 9.0.$$

Here the median (second quartile) is $x_4 = 5.4$. The quartiles are then the medians of the first and second halves (both including x_4). Both halves also have an even number of elements this time. So $Q_1 = \frac{3.9+4.3}{2} = 4.1$ and $Q_3 = \frac{6.7+8.9}{2} = 7.8$ are just the medians of the first and second halves. Something similar would happen for $N = 3, 11, 15$, etc. as well.

It would also be possible, of course, for N to be even, while the halves are odd in length – this happens for instance for $N = 6, 10, 14$, etc. In yet other cases, N might be odd but the halves (including the overall median) could be odd in length – this happens for $N = 9, 13, 17$, etc.[6] $\triangle$

The *inter-quartile range*, sometimes abbreviated to IQR, is defined as the difference $Q_3 - Q_1$. The IQR provides another sort of measure of how spread-out a data set is.

EXAMPLE 9.7. Let us consider the two data sets from Example 9.4 again:

$$5.2 \quad 8.9 \quad 2.1 \quad 7.5 \quad 9.3 \quad 4.3 \quad 4.7$$

and

$$6.2 \quad 7.9 \quad 2.1 \quad 6.5 \quad 9.3 \quad 4.8 \quad 5.2.$$

In sorted order, the first of these is

$$2.1 \le 4.3 \le 4.7 \le 5.2 \le 7.5 \le 8.9 \le 9.3$$

so the median is 5.2, and $Q_1 = \frac{4.3+4.7}{2} = 4.5$, $Q_3 = \frac{7.5+8.9}{2} = 8.2$. The IQR is $Q_3 - Q_1 = 8.2 - 4.5 = 3.7$.

Doing the same computations with the second data set,

$$2.1 \le 4.8 \le 5.2 \le 6.2 \le 6.5 \le 7.9 \le 9.3,$$

so the median is 6.2, and $Q_1 = \frac{4.8+5.2}{2} = 5.0$, $Q_3 = \frac{6.5+7.9}{2} = 7.2$. The IQR is $Q_3 - Q_1 = 7.2 - 5.0 = 2.2$. Recall that we used the standard deviation before and saw that the SD of the first was larger while the SD of the second is smaller. The inter-quartile range indicates that the second data set is less spread-out around its middle in a different way. $\triangle$

---

[6]Technical note: Readers familiar with the notion of integer congruences might recognize that the four cases identified here correspond exactly with the four different possible values $N \equiv 0, 1, 2,$ or $3 \bmod 4$.

Just as the median is less sensitive to small changes in data values than the mean, the inter-quartile range is less sensitive to such changes than the standard deviation. It therefore can be thought of as a *more robust* alternative to the standard deviation that might be more useful in situations where there is uncertainty or variability in the data values being analyzed. However, it should be kept in mind that the SD and the inter-quartile range values are not directly comparable, given what we have discussed so far. For instance, the values of the SD and the IQR values from Examples 9.4 and 9.7 are not especially close to each other.

A Measure of Skewness. The quartiles and the IQR of a dataset can be used to provide a quantitative measure of skewness of a data set. The idea is that *positive skewness* should mean that (Q_3 − median) is greater than (median − Q_1), while *negative skewness* should mean that the reverse is true. In other words, a dataset is positively skewed if the larger values are (*consistently*) farther away from the median (to the right along a number line) than the smaller values are from the median (to the left). A measure that is commonly used is

$$\boxed{\text{skewness} = \frac{(Q_3 - \text{median}) - (\text{median} - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2\cdot\text{median} + Q_1}{Q_3 - Q_1}.} \tag{9.7}$$

EXAMPLE 9.8. Let's return to one of the data sets considered in Examples 9.1 and 9.2:

$$2.1 \le 4.3 \le 4.7 \le 5.2 \le 7.5 \le 8.9 \le 9.3,$$

has mean $\overline{x} = 6.0$, median $= 5.2$, $Q_1 = 4.5$ and $Q_3 = 8.2$. So we compute

$$\text{skewness} = \frac{8.2 - 2\cdot 5.2 + 4.5}{8.2 - 4.5} \doteq 0.62.$$

The fact that the skewness is positive agrees with the observation that the large values are farther away from the median than the smaller ones are.

We also looked at what happens when we remove the value 4.7: The shortened data set

$$2.1 \le 4.3 \le 5.2 \le 7.5 \le 8.9 \le 9.3$$

has mean $\overline{x} = 6.2$, median $= 6.35$, $Q_1 = 4.3$ and $Q_3 = 8.9$. It is often claimed that we should expect a negative skewness when the mean is less than the median. However, note that in this case

$$\text{skewness} = \frac{8.9 - 2\cdot 6.35 + 4.3}{8.9 - 4.3} \doteq 0.11$$

is also positive. The point of this example is that the notion of skewness of a distribution of data values is not completely intuitive. Moreover, the relationship between skewness and the relative positions of the mean and median is not always as direct as saying: "skewed towards larger values ("positive skew") when the mean is greater than median and skewed towards smaller values ("negative skew") when the reverse is true."[7] △

[7] Despite what many elementary statistics textbooks claim!

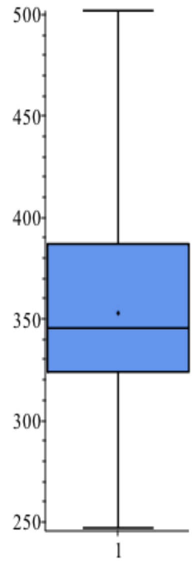


FIGURE 9.1. "Box and whisker plot" of trout length data from Table 3, Chapter 5.

Another Criterion for Outliers. The inter-quartile range is used in another criterion for detecting outliers in a data set originally proposed by John Tukey. Tukey's criterion is that data values *outside the range*:

$$Q_1 - (1.5) \cdot \mathrm{IQR} \le x \le Q_3 + (1.5) \cdot \mathrm{IQR} \tag{9.8}$$

are considered as outliers, where $\mathrm{IQR} = Q_3 - Q_1$ as before. This gives results comparable in many cases to the 3σ rule discussed above.[8] However, there is a certain amount of arbitrariness in the choice of the multiplier 1.5 here and other choices would be possible.

The "Five Number Summary" and Box Plots. When a statistician wants to begin to understand what information might be present in a set of data values, he or she might begin by computing the mean, the SD, and the five statistics

$$\min, Q_1, \text{median}, Q_3, \max. \tag{9.9}$$

Because there is so much information contained in this collection of statistics, it is often called the *five number summary* of a data set.

The five number summary is also often presented in a visual way using what is called a *box plot*, or *"box and whisker" plot.*

EXAMPLE 9.9. Let us consider the data on trout *lengths* contained in Table 3 in Chapter 5 and used in Exercise (12) in that chapter. (You will deal with the weight data in a similar way in Exercise 2.) The "box and whisker plot" is shown in Figure 9.1.

[8]Technical note: The multiplier value 1.5 gives roughly the same 99% fraction of data values within the interval when the data is *normally distributed*, and that is where this choice came from historically.

To interpret this, this data set has $\overline{x} \doteq 352.9$, $s \doteq 57.0$, and five number summary

$$\min = 247, Q_1 = 323.5, \text{median} = 345.5, Q_3 = 387.25, \max = 502.$$

Using the scale on the vertical axis, we can see that

- The horizontal lines in the "box" part of the plot are located at Q_1, the median, and Q_3. So the height of the box is the IQR $= 63.75$.
- The location of the mean is indicated by the small diamond above the horizontal line at the level of the median. Note that the mean is slightly larger than the median in this case. The location of the mean is not always included in a box plot, but it adds information.
- The "whiskers" extend to the maximum and the minimum values in the data set. So the range is just the total height of the plot: $502 - 247 = 255$.

Thus we get a rich visual representation of the distribution of the data by considering this box plot. For instance, we can tell that the skewness will be *positive* just by noticing the top portion of the box (the part between the median and Q_3) is taller than the lower part of the box (the part between Q_1 and the median). $\triangle$

One question that might arise in considering this data is whether the maximum and minimum values should be considered outliers, since they are so far outside the range between Q_1 and Q_3. However, note that

$$502 = \max < \overline{x} + 3 \cdot s = 352.9 + 3 \cdot 63.57 \doteq 542.6,$$

so this value is still within 3 standard deviations of the mean. By the 3σ rule, we would not label it as an outlier. Similarly, the minimum is also within 3 standard deviations of the mean:

$$247 = \min > \overline{x} - 3 \cdot s = 352.9 - 3 \cdot 63.57 \doteq 162.2.$$

You will carry out the analysis using Tukey's criterion from (9.8) in Exercise 4 below.

9.4. Exploratory Data Analysis and Excel

In this section we will consider some larger, real-world data sets and perform some initial *exploratory data analysis.* We will also indicate how the statistics introduced previously in this chapter can be computed in Excel spreadsheets.

Example 9.10. *Darlingtonia californica* is a carnivorous pitcher plant that lives in wetlands in mountains in Oregon and California. It feeds mostly on ants, flies, and wasps that enter a structure on the top of the plant called the "pitcher," and that are then trapped. See the structures at the tops of the tubes making up the plants in Figure 9.2.[9] The unlucky insects are eventually digested by secretions generated in the plant's body.

The data set we will use is a collection of measurements made from 25 individual *Darlingtonia* plants, given in Table 1.[10] The second column gives the overall height of the plant, the third gives the pitcher "mouth" diameter, and the fourth gives the pitcher tube diameter. All measurements are in units of millimeters (mm).

[9] Downloaded from `https://en.wikipedia.org/wiki/Darlingtonia_californica`, August 22, 2017.

[10] This is borrowed from Gotelli and Ellison, *A Primer of Ecological Statistics*, p. 217.

FIGURE 9.2. *Darlingtonia californica* plants.

As a first step in analyzing a collection of data like this, we might compute means, SDs, the five number summaries, box plots, etc. for each of the three measurements. The purposes of this might be to try to identify any issues with the data such as obvious outliers, strange values, etc. and to start to get a feel for the distribution of values of each of the measurements.

After entering the data into an Excel spreadsheet in four consecutive columns A through D in rows 2 through 26, we can compute these statistics for the data in column B, for instance, as follows.

- To compute the *mean* for the data in column B by enter a command `= AVERAGE(B2:B26)` in an unoccupied cell.
- To compute the *standard deviation*, use a command `=STDEV(B2:B26)` in an unoccupied cell, etc.
- To compute the *medians*, use a command `=MEDIAN(B2:B26)` in an unoccupied cell.
- To compute the *quartiles*, use commands like `=QUARTILE.INC(B2:B26,1)` for the first quartile; replace the `1` by `3` for the third quartile.[11]
- To compute the *maximum and minimum*, use commands like `=MAX(B2:B26)` and `=MIN(B2:B26)`.
- Unfortunately, the basic functions of Excel do not include a way to generate box plots directly. There are a number of sets of detailed step-by-step directions for "tricking" Excel into creating box plots that are available online, however. Google "Excel box plot" if you are interested in trying to do this. A warning: however you do it, the process is rather involved(!)

[11]The `QUARTILE.INC` command in Excel computes the quartiles by the *inclusive* method we described above. There is another version called `QUARTILE.EXC` that computes the quartiles *excluding* the median value in the two halves of the list of data when N is odd.

TABLE 1. Measurements from individual pitcher plants, *Darlingtonia californica*

Plant number	Height	Mouth	Tube
1	744	34.3	18.6
2	700	34.4	20.9
3	714	28.9	19.7
4	667	32.4	19.5
5	600	29.1	17.5
6	777	33.4	21.1
7	640	34.5	18.6
8	440	29.4	18.4
9	715	39.5	19.7
10	573	33.0	15.8
11	1500	33.8	19.1
12	650	36.3	20.2
13	480	27.0	18.1
14	545	30.3	17.3
15	845	37.3	19.3
16	560	42.1	14.6
17	450	31.2	20.6
18	600	34.6	17.1
19	607	33.5	14.8
20	675	31.4	16.3
21	550	29.4	17.6
22	5.1	0.3	0.1
23	534	30.2	16.5
24	655	35.8	15.7
25	65.5	3.52	1.77

Fortunately, it is also not difficult to create accurate box plots by hand once you have the five number summary.

For the Height measurements, we have $\overline{x} \doteq 611.7$, $s = 265.1$, and the five number summary is min $= 5.1$, $Q_1 = 545$, median $= 607$, $Q_3 = 700$, and max $=$

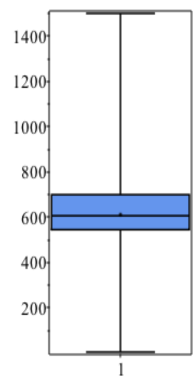


FIGURE 9.3. Box plot of Height data from Table 1

1500. From this we see that the IQR is IQR $= 700 - 545 = 155$. A box plot of the Height data created with a different software package is given in Figure 9.3.

Looking at these results and the values in the Height column from the table, it is pretty clear that there are at least potential issues with the measurements for plants 11, 22, and 25. The Tukey criterion for outliers from (9.8), for instance, says that we should consider values outside the interval

$$312.5 = 545 - (1.5) \cdot (155) \leq \text{Height} \leq 700 + (1.5) \cdot (155) = 932.5$$

as outliers. The values $1500, 5.1, 65.5$ *all lie far outside that interval.* So by the Tukey criterion we would label all three of those measurements as outliers.

On the other hand, the 3σ rule yields more ambiguous conclusions because the interval of values within three SDs of the mean is

$$-183.6 = 611.7 - (3) \cdot (265.1) \leq \text{Height} \leq 611.7 + (3) \cdot (265.1) = 1407.$$

The small values $5.1, 65.5$ are included in this interval, but the maximum value 1500 is excluded. This is actually a very good illustration of a limitation of the 3σ rule. Namely, when there are values in the data set that are very far from the mean and the other data values, the computed s can also be inflated and this makes the interval of values within 3 SDs of the mean very large. Look back at the formula (9.6) to see how a few very large $(x_i - \overline{x})$ terms could produce a large SD value. In particular, in this case, that interval includes a whole range of *negative values* that are *impossible* as Height measurements. So the 3σ rule would flag only the one very large value as an outlier in this data set.

As we said before, though, either way, identifying a data point as an outlier *does not* mean that we will simply delete or ignore it. In the context of the pitcher plants studied here, a height of 1500 mm is not entirely out of the question because some long-lived individuals can definitely grow to more than a meter in height. Similarly, immature plants can be very small, so a height of 66.5 or even 5.1 mm is not impossible either. It is also good to *resist the temptation* to say whoever recorded the data for plants 22 and 25 “clearly” just misplaced the decimal points or forgot the measurements were in millimeters, not centimeters or decimeters. We should also resist the temptation to say there’s a typo in the 1500 (an extra 1). Those are plausible explanations for the data, but we have no definitive reason

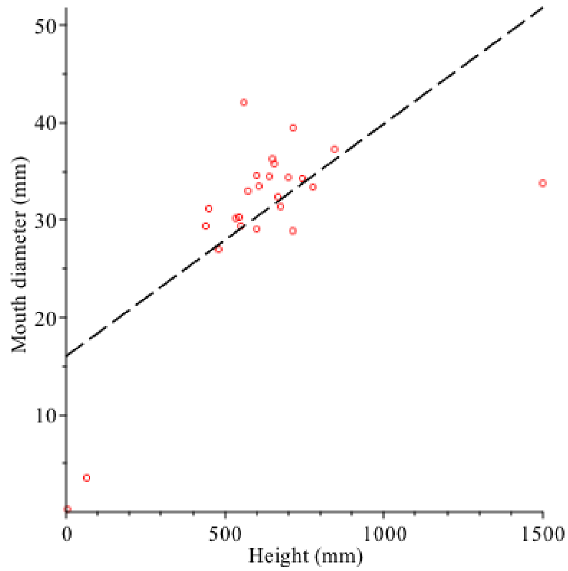


FIGURE 9.4. Scatter plot of Mouth diameter vs. Height data from Table 1

to assert either of those things happened. Editing questionable data to "fix it" according to a theory of how it got that way, however plausible that theory might be, is also considered a very "shady" statistical practice unless you know for certain that an error was introduced in that exact fashion!

You will carry out a similar analysis for the Mouth diameter and Tube diameter data sets in Exercise 7.

Continuing with the exploratory analysis of this data, let's return to the kinds of calculations we did in Chapter 4 on fitting linear models. We can ask, for instance, whether the Mouth diameter of the pitcher plants is linearly dependent on the Height–or, as we said before, whether a linear model is a good fit for this relationship. Figure 9.4 shows the scatter plot of the data points with the Height on the horizontal axis and the Mouth on the vertical axis, and the best-fit regression line is plotted dashed and in black.

The computed R^2 value is $R^2 \doteq .4618$, indicating a rather poor fit for the linear model. The points corresponding to the outliers we identified above are clearly visible here – indeed, the two points very close the origin come from plants 22 and 25, while the point corresponding to plant 11 is the one far to the right of, and below the regression line. We might expect from this picture that the linear relationship would be at least somewhat stronger if the outliers were omitted. But in fact, this is something of an optical illusion because the large ranges of Height and Mouth diameter values in the original data set tend to make the non-linearity of the remaining points less visible. If those three points are omitted, the best-fit regression line changes slightly, but $R^2 \doteq .2011$ is even smaller(!) You will check this in Exercise 9. This sort of omission of some outliers in a data set is called *truncation* in statistics.

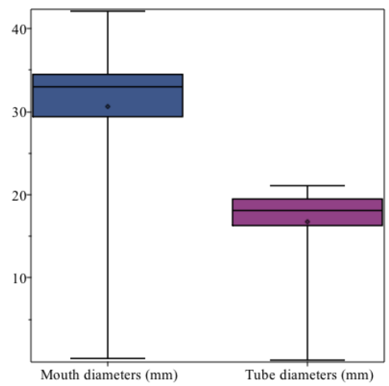


FIGURE 9.5. Parallel box plots for the Mouth diameter and Tube diameter data.

Another exploratory data analysis technique that can be revealing is to plot box plots for two different data sets side by side, or stacked horizontally. This allows one to compare the distributions in a visual fashion. For instance, Figure 9.5 shows the box plots for the Mouth diameter and the Tube diameter data from Table 1, side by side. From this we can see that the Tube diameters are generally smaller, that they also show less variablility (for instance, the IQR is smaller for the Tube diameter data), and both data sets are negatively skewed (i.e. skewed toward smaller values). △

9.5. The R^2 Statistic and the Correlation Coefficient (Optional)

We will conclude this chapter by returning to the computation of the the R^2 statistic that we have used in an informal way to assess the goodness of fit when using least-squares regression, and a related statistic called the *correlation coefficient*.

As we said in (9.1) in the Introduction, this formula:

$$R^2 = \frac{\left(\displaystyle\sum_{i=1}^{N}(x_i - \overline{x}) \cdot (y_i - \overline{y})\right)^2}{\displaystyle\sum_{i=1}^{N}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{N}(y_i - \overline{y})^2}$$

(or an equivalent, rearranged form) is used to compute the R^2 statistic. The correlation coefficient is a square root of this, computed by

$$R = \frac{\displaystyle\sum_{i=1}^{N}(x_i - \overline{x}) \cdot (y_i - \overline{y})}{\sqrt{\displaystyle\sum_{i=1}^{N}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{N}(y_i - \overline{y})^2}}$$

We recognize now that the two factors in the denominator of the formula for R, namely

$$\sqrt{\sum_{i=1}^{N}(x_i - \overline{x})^2} \quad \text{and} \quad \sqrt{\sum_{i=1}^{N}(y_i - \overline{y})^2}$$

multiply to give $(N-1)$ times the product of the SDs of the x_i and the y_i. In the formula for R^2, we have the squares of these. The square s^2 of the SD for a data set is called the *variance*, so the denominator in the formula for R^2 is $(N-1)^2$ times the product

$$(\text{variance of the } x_i's) \cdot (\text{variance of the } y_i's).$$

The expression

$$\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x}) \cdot (y_i - \overline{y}),$$

related to the numerator in the formula for R^2, is known as the *covariance* of the the x_i and the y_i. If the x_i and the y_i were equal for all i, then the covariance would reduce to the variance of that single data set. When the two data sets are different, the covariance, and the correlation coefficient are designed to measure the degree to which the y_i *depend linearly on* the x_i (or vice versa). We can see the connection if we look at the case where $y_i = mx_i + b$ for some constant m, b and all i – or equivalently, where the data points (x_i, y_i) all lie exactly on some straight line. If that is true, then

$$(y_i - \overline{y}) = (mx_i + b - (m\overline{x} + b) = m(x_i - \overline{x}).$$

Substituting this into the formula for R and simplifying, we have (when there is an exact linear relation between the y_i and the x_i):

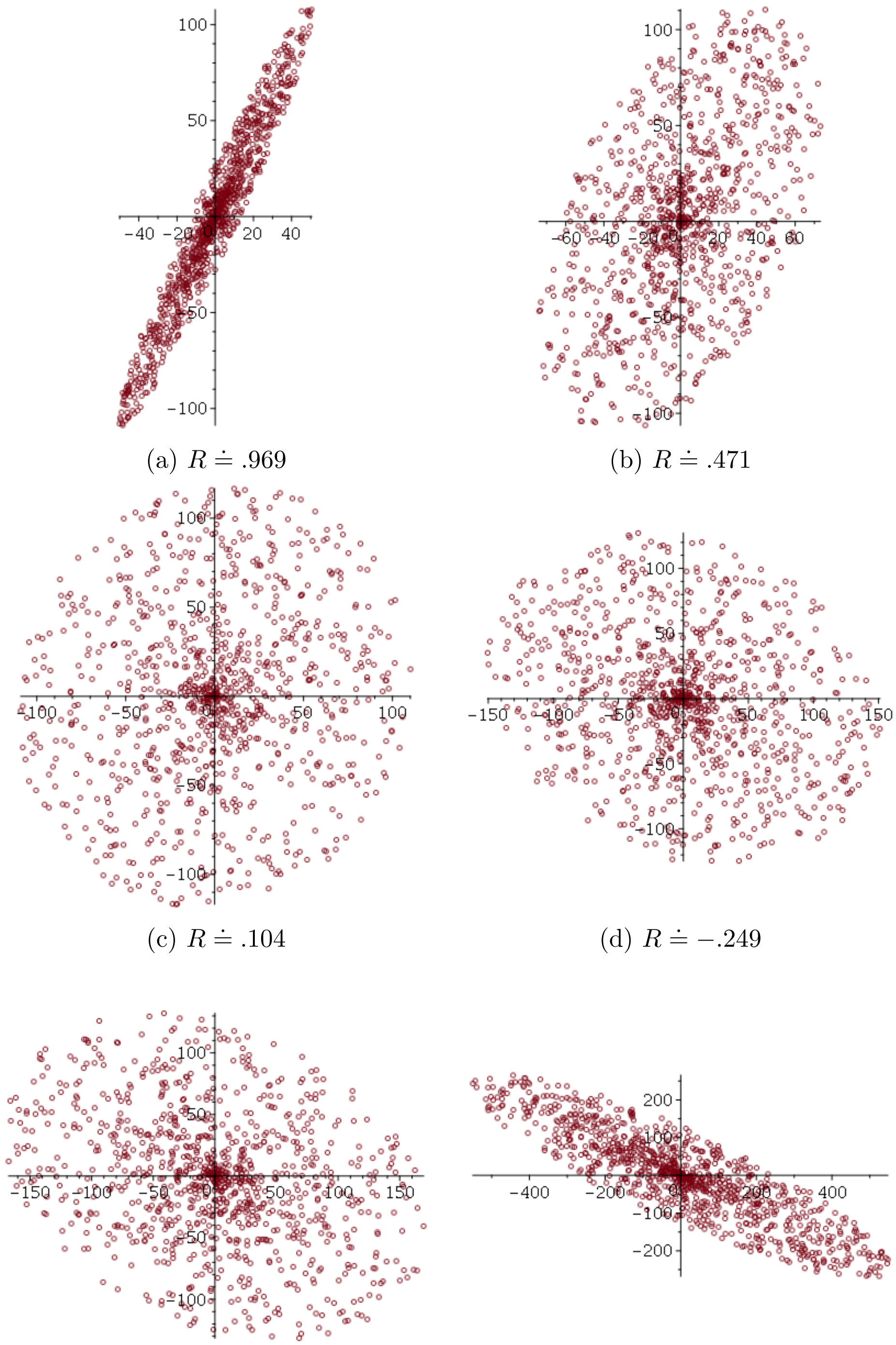
$$R = \frac{m \cdot \sum_{i=1}^{N}(x_i - \overline{x})^2}{\sqrt{m^2} \cdot \sum_{i=1}^{N}(x_i - \overline{x})^2} = \frac{m}{|m|} = \begin{cases} +1 & \text{if } m > 0 \\ -1 & \text{if } m < 0 \end{cases}. \tag{9.10}$$

Hence $R = \pm 1$ in these cases and it can be shown that $-1 \leq R \leq 1$ in all cases. Moreover, the closer R is to the endpoints of the interval from -1 to $+1$, the closer the data points (x_i, y_i) are to lying along a straight line.

We can see this in intuitive terms by considering several data sets as in Figure 9.6. The figure shows six synthetic data sets with $N = 1000$, with R-values ranging from something close to $+1$ in (a) to something close to -1 in (f).[12] We note the following observations:

- As the R values decrease from .969 to .104 in subfigures (a), (b), (c), the "point cloud" gets progressively less linear. In subfigure (c), in fact, the cloud is almost circular in shape with essentially no linear relationship between the x_i and y_i.
- The regression line would still have positive slope in all of the first three cases, however.
- The slope of the regression line turns negative in subfigures (d), (e), (f).
- As the R values decrease further from $-.249$ to $-.890$, the "point cloud" gets progressively more linear again.

[12]Technical note: These were generated from *bivariate normal* distributions where the correlation coefficient can be adjusted as a parameter.



(a) $R \doteq .969$ (b) $R \doteq .471$

(c) $R \doteq .104$ (d) $R \doteq -.249$

(e) $R \doteq -.307$ (f) $R \doteq -.890$

FIGURE 9.6. The meaning of the correlation coefficient

This is a somewhat more precise indication of the background behind our use of the R^2 statistic in regression analysis. We will discuss more of the geometry of regression and making inferences about regression coefficients in a later chapter.

9.6. Chapter Project

The descriptive statistics we have introduced in this chapter are important tools, but they are not the end of the story. The basic goal of statistics is to analyze the information that is contained in a data set and to make inferences from that information. The data sets used in environmental questions might come from sampling some larger population of organisms. Even when that is not the case, we can use the idea of sampling from a larger population of all possible measurements as a way to conceptualize how the data was generated.

To understand some of the statistical topics we will look at starting in the next chapter, it will be good to have some more intuitive understanding of the process of taking samples (random or otherwise) from a population, and the variability that is built into the process of computing statistics that are produced by that sampling process. This may seem counterintuitive at first – after all, there is just one mean or SD (for instance) associated to any particular sample from a larger population. There is no variation or uncertainty in dealing with any one data set. But the idea is that those sample means and SDs and other statistics produced from *different* sample data sets *can vary a lot* and we need to have a good working understanding of several aspects of that sort of variation.

We will also practice with some additional features of Excel that may be useful for your work on larger assignments for this and other courses. Unlike some of the other chapter projects, this project will not focus on an environmental data set, but instead will use synthetic data produced in different ways. Most of your work for this will be contained in a spreadsheet file.

Constructing a First "Population" and Sampling From It. Open a new Excel spreadsheet and enter the formula `=RAND()` in cell A1. This function computes a random value from a uniform distribution on the interval 0 to 1.[13] Copy and paste that into all the cells in the block extending from column A through column J and row 1 through row 20 (200 cells in all). Show 3 decimal places in all of the numbers. You now have a random sample of size 200 from that uniform distribution.

Because of a "quirk" of the way Excel deals with spreadsheets containing this `RAND()` function in formulas, you will see that if you now enter a formula in another cell of the spreadsheet and calculate that other value, all of the numbers in this block will be recomputed too. To turn that feature off (it's annoying and it will mean your results are constantly changing!) highlight the whole block and from the tool bar select Edit/Copy. Then Edit/PasteSpecial, highlight only the option for Values (or Numbers in Libre Office), then OK. If you examine the contents of the

[13]Technical note: The outputs of `RAND` are called *random numbers*, but they are produced using a definite algorithm, as is true for essentially everything done with computer software. So a better name would be *pseudo-random numbers*. They are only random in the sense that they have no apparent patterns and if you compute a large collection of them, their distribution will be approximately uniform over the interval from 0 to 1.

cells in the block now, you should see just numbers, not the formulas that generated them.

First Calculations. Now perform the following calculations:

(A) Compute the average (mean) and the SD of the 200 numbers. Compute the five-number summary and construct a box plot (by hand or with Excel, if you are feeling adventurous and confident).

(B) Next compute the data needed for a *frequency histogram* showing the distribution of this population of 200 numbers (look back at Chapter 6 if you need a refresher on this). Here's a "shortcut" Excel method for doing this. Since all the numbers are between 0 and 1 and we have a relatively large data set with $N = 200$, we can use 10 equal bins with upper boundaries at $0.1, 0.2, \ldots, 1.0$. Enter those numbers into cells in one column in the spreadsheet at least three rows below the block of 200 cells above (say in column `A` and rows 31 to 40). Then in the parallel column `B`, highlight the squares in all 10 rows next to the bin boundaries (cells `B31:B40`). In the last cell at the bottom enter the command: `=FREQUENCY(A1:J20,A31:A41)`, and press `SHIFT-CTRL-ENTER` (all together) (or `F2` then `SHIFT-CTRL-RETURN` on a Mac) all together. The whole frequency table should be computed and displayed in the selected cells.

(C) Generate a frequency histogram using the computed frequencies with the bins as above.

(D) Now, we consider sampling from this "population." Notice that the block `A1:J20` can be thought of as either:
- 20 rows (horizontal) with 10 cells in each, or as
- 10 columns (vertical) of 20 cells each.

We will think of those as two different groups of samples from the population consisting of all 200 of the numbers–one group of 10 samples of size 20 from the columns, and as second group of 20 samples of size 10 from the rows. (These are not really random samples, of course, since they are constructed in this systematic way. However the whole block was computed by processes that produce random-looking results, so it will do no harm to think of them as random.)

(1) Compute the averages of each of the 20 rows and put those averages in the cells `L1:L20` (i.e. in the cells in column `L` parallel to the block)

(2) Find the mean and SD of the 20 row averages from part 1. Generate a frequency histogram for those 20 values using 5 bins. (You will need to decide on appropriate bin boundaries by examining the values you get.)

(3) Compute the averages of each of the 10 columns and put those averages in the cells `A22:J22` (i.e. in the cells in row 22 under the block).

(4) Find the mean and SD of the 10 column averages from step 3. Since we only have 10 of these, making a frequency histogram is not too meaningful, so you do not need to do that.

Constructing a Second Population and Sampling From It. Open a second tab (sheet) in your spreadsheet and enter the formula

```
=SQRT(-2*LN(1-RAND()))*COS(RAND()*2*PI())
```

in cell `A1`. (This might be tricky – if you get an error message, check the parentheses carefully and make sure they are exactly as above.) Copy and paste that into all the cells in the block extending from column `A` through column `J` and row 1 through row 20 (200 cells in all). Show 3 decimal places in all of the numbers. Repeat parts A – D above for this new data set, except that you will need to decide on appropriate bin boundaries for the histograms. (Don't worry about exactly what the formula means. The only thing you need to know for now is that it's designed to produce different random values from the ones you used in the first section.)

Further Questions.

(E) Note that `RAND()` appears in both formulas you used to generate both of the "population data sets? above. Again, this is a built-in Excel formula that generates random (more precisely "pseudo-random") numbers in the interval from 0 to 1. Looking at your population histogram from the first data set, what can you say about the distribution of those numbers? Do they seem to be reasonably uniformly distributed? What about the second data set where you used the ?black magic? formula? How is that one different? How do the two population box plots compare?

(F) What about the distributions of the row and column averages in both cases: what did the distributions look like there? Did it depend which data set they came from?

(G) What can you say about the SD's of the original population, the SD's of the averages of the samples of size 20, and the SD's of the averages of the samples of size 10? In particular, was one of these consistently the largest and one consistently the smallest? If the pattern was not entirely consistent, was there a general trend?

(H) If we redid these computations with any fixed rule for computing the "population" values, would it always be true that the average of the row means is the same as the population mean? Similarly for the average of the column means? (Hint: algebra!)

(I) What would truly random samples of sizes 10 or 20 from these populations look like? Suggest a way to generate such samples.

(J) If we had 20 truly random samples of size 10 from either of these populations, would the average of the sample means be the same as the population mean? (This is the same question as H above, but for a different way of producing the samples.) Explain your answer.

Assignment. Submit your spreadsheet and your answers to questions E - J in a separate document.

Chapter Exercises

(1) Compute (by hand) the mean, median, range, and standard deviation of the following hypothetical data set:

$$4.8 \quad 2.9 \quad 5.4 \quad 4.3 \quad 5.1 \quad 3.2 \quad 39 \quad 4.7.$$

(2) Using Excel, compute the mean, standard deviation, and five number summary for the fish *weight* data from Table 3 in Chapter 5 used in Exercise (12) in that chapter. (The *length* data is analyzed in the text.)

(3) Using Excel, compute the mean, standard deviation, and the five number summary for the tanker oil spill amounts from Table 5 at the end of Chapter 6.

(a) How many of the spills are within one standard deviation of the mean?

(b) Is this data set positively or negatively skewed? Draw a box plot to verify your answer.

(4) Are the maximum and minimum values in Example 9.9 outliers according to Tukey's Criterion (9.8)? Why or why not?

(5) This problem gives an alternate formula for the standard deviation.

(a) Show that the formula given in (9.6) in the text can be rearranged by means of algebra to yield

$$s = \sqrt{\frac{1}{N-1}\left(\sum_{i=1}^{N} x_i^2 - N \cdot \overline{x}^2\right)}.$$

(Hint: Expand out the square in each term in the form given in the text. Then simplify using the formula for $\overline{x}$.)

(b) Use this alternate form to compute the standard deviations of the data sets from Example 9.4. You should find that the computations are somewhat simpler because you do not need to compute the deviations. You can work just with the x_i and with $\overline{x}$.

(6) Show that the formula for the R^2 statistic can be rearranged to each of the following forms:

$$R^2 = \frac{N \cdot \sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i}{\sqrt{N \cdot \sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)^2} \cdot \sqrt{N \cdot \sum_{i=1}^{N} y_i^2 - (\sum_{i=1}^{N} y_i)^2}},$$

which also simplifies to

$$R^2 = \frac{\sum_{i=1}^{N} x_i y_i - N \cdot \overline{x} \cdot \overline{y}}{\sqrt{\sum_{i=1}^{N} x_i^2 - N \cdot \overline{x}^2} \cdot \sqrt{\sum_{i=1}^{N} y_i^2 - N \cdot \overline{y}^2}}.$$

(7) Carry out the sort of analysis given in Example 9.10 for the Mouth diameter and the Tube diameter data sets in Table 1.

(a) Compute the mean, SD, and five number summary of each data set. Construct a box plot for each by hand, or using Excel if you are feeling adventurous(!)

(b) Are there any outliers in those data sets?

(c) Does there appear to be a linear relation between the Tube diameter and the Height?

(8) Show that the three datasets in Table 1 (the Height, the Mouth diameter, and the Tube diameter) all satisfy the statement of Chebyshev's theorem with $n = 2$.

(9) Verify the claims made at the end of Example 9.10 regarding the effects of truncation of outliers in the Mouth diameter versus Height bivariate dataset. In particular, generate the point plot of the truncated dataset, plot the regression line, and show that the R^2 value decreases to $R^2 \doteq .2011$.

(10) Investigate the relation between the Mouth diameter and the Tube diameter in Table 1.

(a) Does there seem to be a linear relationship there?

(b) What happens if the data is truncated to remove outliers? Does the R^2 statistic move closer to 1?

(11) The data in Table 2 gives the areas and the numbers of species of plants inhabiting each of the small islands making up the Galapagos chain off the coast of South America.[14]

TABLE 2. Galapagos Island Plant Species Diversity

Island	Area (square km)	Number of Species
Isabela	5824.9	325
Floreana	165.8	319
San Cristobal	505.1	306
Santiago	525.8	224
Santa Cruz	1007.5	193
Pinta	51.8	119
Pinzon	18.4	103
Fernandina	634.6	80
Espanola	46.6	79
Seymour	2.6	52
Santa Fe	19.4	48
Gardner	0.5	48
Marchena	116.6	47
Rabida	4.8	42
Genovesa	11.4	22
Wolf	4.7	14
Darwin	2.3	7

[14]This was one of the stops of the H.M.S. *Beagle* where Charles Darwin collected samples that he eventually used as examples to support his theory of evolution by natural selection. Data from F.W. Preston, The canonical distribution of commonness and rarity: Part I, Ecology **43** (1962), 185–215.

(a) Compute the mean, SD, and five number summary for both the Area and Number of Species datasets.

(b) Are there outliers according the to the criteria discussed in the text? Would it be reasonable to omit those data items if there are? Explain.

(c) Investigate the goodness of fit of linear, exponential, and power law models for the Number of Species as a function of the Area of the islands. Which type of model seems to fit best?

CHAPTER 10

Random Variables and Probability Distributions

10.1. Introduction

In Chapter 2, we discussed how ratios are used to quantify the probability of an event in the standard, frequentist developments of probability theory. Recall that the idea was that if we perform some (possibly large) number of trials of an experiment or measurement, then the probability of a certain event or outcome could be estimated by a ratio of the form:

$$\frac{\text{Number of times event is observed}}{\text{Total number of trials}}. \tag{10.1}$$

For instance, if we throw a fair standard six-sided die some number of times, we could count the number of times each of the six possible numbers comes up and divide each of those counts by the total number of throws. In this situation, we would usually also have a theoretical *probability model* of a fair die in mind where the probability of each number coming up is exactly $\frac{1}{6}$. While we expect some deviation from the exact $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$ distribution in any real-world instance of this sort of experiment, if the die is not "loaded" to a significant degree, we would also expect the ratios as in (10.1) to be close to $\frac{1}{6}$ for each of the 6 numbers.

In this chapter we will develop more of the theory behind probability distributions in both discrete and continuous cases. In that sense, this chapter and the next one will be somewhat different from the others, since the goal here will be to develop some new mathematical and statistical tools, and not so much to show how those tools used for environmental questions.

The starting point will be a discussion of several additional concepts from the theory of probabilities, culminating in the important notion of *independence* of events. We will introduce several types of discrete and continuous random variables and discuss computing their expected value and variance. Finally, we will introduce an important family of continuous probability distributions known as the *normal distributions*. In the Chapter Project, we will illustrate a key result known as the Central Limit Theorem, which shows that normal distributions naturally arise in processes where we average large numbers of measurements. The formal statement will be given in a later chapter.

10.2. Conditional Probabilities and Independence

In section 4 of Chapter 2, we discussed assigning probabilities to events E in sample spaces S in an intuitive way. More formally, we can also describe the properties that such an assignment of probabilities should have as in Assumptions 10.1

below. In this and later in this section, we will need to use the following notation from basic set theory. If A, B are sets, then

- $A \cup B$ is the *union* of A and B, defined as the set consisting of all elements that are *in A, in B, or in both.*
- $A \cap B$ is the *intersection* of A and B, defined as the set consisting of all elements that are *both in A and in B.*

For instance if $A = \{1, 2, 3, 4, 5, 6, 7\}$ and $B = \{2, 4, 6, 8, 10\}$, then

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 10\}$$

and

$$A \cap B = \{2, 4, 6\}.$$

Assumptions 10.1. If S is a sample space for an experiment we will assume we have a way to assign a probability $P(A)$ to each subset $A \subseteq S$ satisfying the following properties

(1) $P(A) \geq 0$ for all $A \subseteq S$,
(2) $P(S) = 1$, and
(3) If $A_1, \ldots, A_n$ is any sequence of *pairwise disjoint* subsets of S[1] then

$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n).$$

The last assumption here, for instance, is what is needed to justify computations of probabilities of compound events by summing probabilities of their elements as in Examples 2.11 and 2.12 from Chapter 2.

Conditional Probabilities. Suppose we know that we have rolled an odd number on roll of a single fair die. What is the probability that the roll is a 1? A moment's thought will convince you that the answer here should be 1/3. Knowing that the roll is odd cuts the possibilities down to $B = \{1, 3, 5\}$ and if the die is fair, each of those three possibilities is equally likely. This is an example of what is known as a *conditional probability.* The formal definition looks like this:

Definition 10.2. The *conditional probability* of an event A, given that another event B with $P(B) \neq 0$ has occurred,[2] written $P(A|B)$, is defined to be

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}.} \tag{10.2}$$

In other words, we in effect restrict our attention to a new sample space consisting of the subset $B \subseteq S$ and consider the probability of A in that new context, ignoring what happens for any part of A outside of B. For future reference, we note that (10.2) can also be rearranged to the form

$$P(A \cap B) = P(A|B)P(B). \tag{10.3}$$

Example 10.3. The simple example given at the start of this subsection translates into this language as follows. The set of odd rolls $B = \{1, 3, 5\}$ sits inside the sample space $S = \{1, 2, 3, 4, 5, 6\}$ for the roll of a single die. Since the die is fair,

[1] This means that $A_i \cap A_j = \emptyset$, the empty set, whenever i and j are distinct indices.

[2] Saying an event "has occurred" is too suggestive to resist, but it really just means that we are looking at outcomes in the intersection $A \cap B$.

each roll has probability 1/6, so $P(B) = 3/6 = 1/2$. $A = \{1\}$ so $A \cap B = \{1\}$ as well. Hence

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = 1/3,$$

as we expected.

Similarly, if we throw two distinguishable fair dice (say one with white background and dark spots, and the other with dark background and white spots) as in Example 2.12, we can ask, given that the total on the dice is a 7, what is the probability that one of the dice shows a 5 or a 6? Here B represents the possible throws yielding a total of 7:

$$B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}.$$

A is the event that one of the two dice shows a 5 or a 6. That is true in 4 of the 6 elements in B, so we can see $P(A|B) = 4/6 = 2/3$. This would also come from the definition, since

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{4/36}{6/36} = 2/3.$$

On the other hand, you will verify in Exercise 1 that $P(A) = 5/9 < 2/3$. $\triangle$

In both cases considered in Example 10.3, it turned out that $P(A|B)$ was greater than $P(A)$. It is also possible for $P(A|B)$ to be equal to $P(A)$ or less than $P(A)$. For example, if $A \cap B$ is empty, then $P(A|B)$ will usually equal zero, even if $P(A) \neq 0$. In any case the idea is that knowing B *gives more information* and hence can change the probability of A, making it more or less likely in some circumstances. Here is a famous and counterintuitive example of this phenomenon.

Example 10.4. On the classic daytime TV show "The Price is Right," the host Monty Hall would often play the following game with contestants. Monty would present three doors. Behind one was a desirable prize, a new car, while what were considered undesirable "prizes," live goats, were behind the other two.[3] The contestant would be asked to choose one of the three doors. Monty would then reveal one of the two remaining doors with one of the goats. Note that whichever door the contestant chose to begin with, there would be a goat behind one of the two remaining doors, and *very importantly*, Monty always knew which doors had which prizes. At this point, the contestant was asked whether he or she wanted to stick with the door he or she chose to start, or whether he or she wanted to switch to the other remaining door.

The question is: What was the best strategy for the contestant to use–stick with the original choice, or switch?

If you have never heard of this example before, you are encouraged to pause and experiment by simulating this game with three playing cards and a friend before continuing to the next example. Take an ace (representing the car) and two twos (representing the goats). Have your friend play Monty Hall, who always knows where the goats are and shows you one of them, giving you the chance to switch. Play the game 25 or 50 times using the "always stick with the original choice" strategy and record how many times you win the car. Then repeat using

[3]There are certainly some people in the world who would in fact be happier to win a live goat than a car, so the game probably only works as intended for American or other "first-world" contestants!

the "always switch" strategy and record how many times you win the car in 25 or 50 tries. △

EXAMPLE 10.5. If you carried out the simulation suggested in Example 10.4, you very likely found that always switching yielded significantly better results than always sticking with your original choice. Here is the explanation.

If you stick with your original choice, then everything after that first choice, including Monty revealing one of the goats, is actually *irrelevant*(!) Your probability of winning the car is exactly $1/3$ because only one of the three doors has the car.

On the other hand, if you always switch, then something very interesting happens. If you chose the door with the car to begin with, then you will lose because you switch to the door with the other goat (the one Monty did not show). *However*, if the door you chose originally had a goat, then *you always win the car by switching* because Monty showed you the other goat and the remaining door must be the one with the car. Moreover, since there were two goats out of the three choices, the probability that you chose a goat to begin with was $2/3$(!) We can say this formally using the language of conditional probabilities and (10.3) as follows. Let A be the event that you win the car and B_1 be the event that you choose the car to start, while B_2 is the event that you choose one of the two goats to start. Note that $B_1 \cup B_2 = S$, the whole sample space for your original choice. Moreover $B_1 \cap B_2 = \emptyset$. By the result of Exercise 2, if you stick to your original choice, then

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) \\ &= 1 \cdot (1/3) + 0 \cdot (2/3) \\ &= 1/3. \end{aligned}$$

On the other hand, under the "always switch" strategy,

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) \\ &= 0 \cdot (1/3) + 1 \cdot (2/3) \\ &= 2/3. \end{aligned}$$

An interesting sidelight to this example is that essentially this analysis was publicized by Marilyn Vos Savant in her *Parade* magazine column "Ask Marilyn" in several articles in 1990 and 1991. A number of professional mathematicians wrote her to chide her for missing the "obvious" observation that there should be a $50-50$ chance of winning the game after Monty revealed the first goat–switching or not should make no difference. *They were wrong!* Conditional probabilities can be tricky and counterintuitive. △

The computations we did in Example 10.5 are special cases of an important result called the *Law of Total Probability.* Suppose we have a sample space S that is the union of pairwise disjoint subsets $S = S_1 \cup \cdots \cup S_n$. If A is any event in S, then by Assumptions 10.1 and (10.3),

$$\begin{aligned} P(A) &= P(A \cap S_1) + \cdots + P(A \cap S_n) \\ &= P(A|S_1)P(S_1) + \cdots + P(A|S_n)P(S_n). \end{aligned}$$

Independence. We now come to a central concept in probability, namely what it means to say that two events are *independent*. The correct intuition here is that an event A and an event B are independent exactly when knowing that B occurred *has no effect on the probability of* A. In the form of an equation, we say A, B are independent exactly when

$$P(A|B) = P(A).$$

Since $P(A|B) = \frac{P(A \cap B)}{P(B)}$ from (10.2), this condition can also be rewritten in the following form

$$\boxed{P(A \cap B) = P(A)P(B).} \tag{10.4}$$

From this we see that if $P(A) \neq 0$, then by dividing by $P(A)$ the independence relation has a symmetric nature:

$$P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B) \Leftrightarrow P(A \cap B) = P(A)P(B).$$

Note that saying A and B are independent is *almost always different from* saying that $A \cap B = \emptyset$. Here is a simple example.

Example 10.6. Consider one throw of the two distinguishable fair dice from Example 10.3 above. Let A be the event that the first die shows an even number and let B be the event that the second die shows an odd number. We claim that A and B are independent. Since the even numbers are half of the possibilities for the first die while the second die can show any number in A, we have $P(A) = 18/36 = 1/2$. Similarly, $P(B) = 1/2$. The event $A \cap B$ consists of the pairs

$$\{(2,1),(2,3),(2,5),(4,1),(4,3),(4,5),(6,1),(6,3),(6,5)\}.$$

Hence $P(A \cap B) = 9/36 = 1/4 = P(A)P(B)$. This shows that A and B are independent. $\triangle$.

10.3. Discrete Distributions and Probability Mass Functions

A random variable can be thought of as a *function defined on a sample space.* We will begin by studying several commonly-encountered *discrete* random variables and their probability distributions. In these cases, the sample space is finite, or perhaps countably infinite.[4]

Binomial and Hypergeometric Random Variables. One of the simplest examples begins with a sample space consisting of just two elements, often denoted 1 and 0 or "success" and "failure."[5] A *Bernoulli trial* is a random variable X that models the process of performing an experiment with two possible outcomes, each occurring with *fixed probabilities.* The probability of the "success" outcome is

[4]Technical Note: A countably infinite set is one that can be put into one-to-one correspondence with the natural numbers, $\mathbb{N} = \{1, 2, 3, \ldots\}$. A celebrated result of the 19th century German mathematician Georg Cantor shows that there are also infinite sets that are not countably infinite. These are known as *uncountably infinite* sets. In particular, the interval $[0, 1]$ in the real numbers is uncountably infinite. The proof that establishes this result is sketched in the (optional) Exercise 17 below.

[5]The "success" and "failure" language is merely a convention. You should not necessarily associate a "success" in this sense with a "good" outcome or a "failure" with a "bad" one.

$P(X = 1) = p$ for some real number $0 \le p \le 1$, and then the probability of the other outcome is $P(X = 0) = 1 - p$. This complementary "failure" probability is often denoted by $q = 1 - p$.

We use Bernoulli trials to define what is known as a *binomial experiment.* A binomial experiment is one in which:

(1) there are a fixed number, n, of Bernoulli trials, all with the same success probability p,
(2) the trials are *independent* in the sense of (10.4)–outcome of one trial has no effect on the outcomes of the others.
(3) the random variable of interest is $X =$ the *number of successes.*

If these conditions are satisfied,[6] then we say X is a *binomial random variable.*

EXAMPLE 10.7. A standard example of a binomial experiment given in many traditional textbooks on probability theory consists of the following: Imagine we have a large urn filled with balls of two colors – say red and black. Each trial consists of drawing one ball at random from the urn and recording its color.[7] After each trial, the ball is *replaced in the urn* and the contents are mixed thoroughly. At the end, the number of red balls chosen in the n trials is totaled.

Because the ball is replaced after each trial, the numbers of balls of each color in the urn stay the same on each of the trials. If there are n_1 red balls and n_2 black balls, and we associate choosing red ball with the "success" outcome, this means that the probability p of choosing a red ball and the complementary probability of choosing a black ball will be

$$p = \frac{n_1}{n_1 + n_2} \qquad \text{and} \qquad q = \frac{n_2}{n_1 + n_2}$$

on each trial.

The *independence* referred to above consists in the fact that the outcome of any one of the trials should have no influence on the outcomes of the others. This is ensured by the fact that the ball selected on any one trial is coming from the same collection of balls and the random selection should mean that the previous choices have no influence on the later choices. In more precise terms, what this means is that the probabilities of particular outcomes on the successive trials are *multiplied* to obtain the probability of a sequence of trials. That is, the probability of drawing red, black, red, red on four successive draws is

$$p \cdot q \cdot p \cdot p = p^3 q$$

under the assumption that the trials are independent (using (10.4)).

The random variable of interest in a binomial experiment is the number of successes. However, different sequences of draws will lead to a given number of successes in most cases. For instance, with $n = 4$ trials we could end up with three red balls with any of the following sequences of outcomes of the individual trials:

$$\begin{array}{l} \text{red, red, red, black} \\ \text{red, red, black, red} \\ \text{red, black, red, red} \\ \text{black, red, red, red} \end{array}$$

[6] possibly after interpreting a problem to put it into this framework

[7] You are supposed to visualize reaching into the urn to select a ball without being able to see any of the colors before the ball is pulled out!

These are distinct, disjoint outcomes so the total probability of getting $k = 3$ reds in drawing $n = 4$ balls from the urn would be

$$p^3q + p^3q + p^3q + p^3q = 4p^3q.$$

We can think of the coefficient 4 here as the total number of ways to distribute the three reds over the four trials. △

In general, now, if we have a binomial experiment with n trials, the success probability $= p$, and we want the probability of k successes, then we need to have a way to count the number of ways to distribute the k successes over the n trials. It is not difficult to see that this is *the same as the coefficient of* p^kq^{n-k} *in the binomial expansion*:

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^kq^{n-k}.$$

The reason, of course, is that in expanding out the power $(p+q)^n$, after rearranging the products to put p's before q's, the terms of the form p^kq^{n-k} will come from exactly the same sort of selection of k of the p terms in the factors and $n-k$ of the q terms. The number $\binom{n}{k}$ multiplying the p^kq^{n-k} term in the expansion is called the *binomial coefficient* and is often read as "n choose k."[8] Fortunately, there is a convenient formula for $\binom{n}{k}$ using the idea of the *factorial* of an integer. Recall that if n is a positive integer

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1.$$

So, for example, we have

$$\begin{aligned}
1! &= 1\\
2! &= 2 \cdot 1 = 2\\
3! &= 3 \cdot 2 \cdot 1 = 6\\
4! &= 4 \cdot 3! = 24\\
5! &= 5 \cdot 4! = 120
\end{aligned}$$

and so forth. By convention, $0! = 1$. The formula for the binomial coefficient is

$$\boxed{\binom{n}{k} = \frac{n!}{k!(n-k)!}}. \tag{10.5}$$

For example $\binom{4}{3} = \frac{4!}{3!1!} = \frac{24}{6} = 4$, as we saw at the end of Example 10.7. Note that we can also cancel the factors in $(n-k)!$ between the numerator and denominator, yielding the following somewhat more efficient formula

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}.$$

Putting everything together, we have what is called the *probability mass function* for a binomial random variable X:

$$\boxed{P(X = k) = \binom{n}{k} p^kq^{n-k}}. \tag{10.6}$$

The values where $P(X = k)$ would be nonzero are $k = 0, 1, \ldots, n$. See Exercise 10 below for some general properties of this and other probability mass functions.

[8]Technical Note: If you have learned about combinations in high school, this is the same as the number of combinations of n things taken k at a time.

EXAMPLE 10.8. A new surgical procedure has a success probability of p each time it is performed. Here each individual surgery done with the procedure constitutes a Bernoulli trial, and under suitable assumptions (what assumptions? – see below) we can apply the binomial probability formulas.

(1) If $p = .8$ and the surgery is performed 5 times, what is the probability that the surgery will succeed for exactly 4 patients? With $p = .8$, $q = 1 - .8 = .2$. This is the binomial probability

$$P(X = 4) = \binom{5}{4}(.8)^4(.2) = .4096.$$

(2) If $p = .7$, what is the probability that the surgery will succeed for at least two patients? With $p = .7$, $q = 1 - .7 = .3$. The direct way to compute that binomial probability would be

$$P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5).$$

This would yield the value

$$\binom{5}{2}(.7)^2(.3)^3 + \binom{5}{3}(.7)^3(.3)^2 + \binom{5}{4}(.7)^4(.3) + \binom{5}{2}(.7)^5.$$

A somewhat easier, but equivalent, computation would be to take the complementary probability that the surgery is successful 0 or 1 times and subtract that from 1:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) \\ &= 1 - (P(X = 0) + P(X = 1)) \\ &= 1 - \left(\binom{5}{0}(.3)^5 + \binom{5}{1}(.7)(.3)^4\right) \\ &\doteq .9692. \end{aligned}$$

We are assuming that the outcomes of different operations are independent in order to apply the binomial probability formulas in this way. $\triangle$

EXAMPLE 10.9. Suppose we have a large tank containing 334 male trout and 275 female trout. We pick 10 of the fish at random one at a time, and replace them in the tank after determining the sex. We ask: what is the probability that exactly 6 of the fish selected were females?

This is exactly like Example 10.7 with the red balls corresponding to the females and the black balls to the males. So it can be modeled as a binomial experiment with $n = 10$, and $p = \frac{275}{275+334} = \frac{275}{609} \doteq .4516$. If X is the number of females selected, using the more efficient cancelled formula we have

$$\binom{10}{6} = \frac{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$$

and then

$$P(X = 6) = \binom{10}{6}(.4516)^6(.5484)^4 = 210 \cdot .008482 \cdot .09045 \doteq .1611.$$

That is, there is about a 16% chance of this happening – four females is the most likely single outcome because of the imbalance in numbers.

We can also represent the probability mass function values for all integers k, $0 \leq k \leq 10$, with a bar chart like the one in Figure 10.1. The bars extend from $k - \frac{1}{2}$ to $k + \frac{1}{2}$ so their widths are equal to 1 and they are centered on the integer

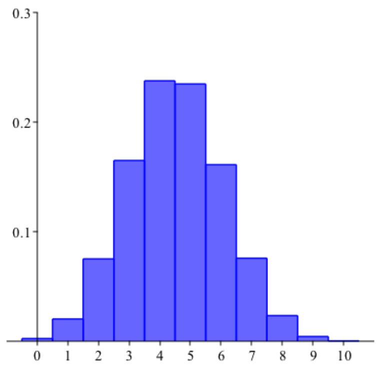


FIGURE 10.1. Binomial probabilities, $n = 10$, $p \doteq .4516$.

k. The sum of the heights of the bars is 1 and hence the total area of the bars sums to 1. This is an example of a *probability histogram*. There is a corresponding probability histogram for each n and each choice of p. When $p \neq .5$, the histogram will not be symmetric. $\triangle$

In Example 10.9, note that we were careful to say that the fish were replaced in the tank after each one was selected. But that probably seems like an unreasonable way to carry out the process of selecting 10 of the fish because we could be selecting the same fish more than once. If we did not do the selection with replacement, then the process would be approximately, but not exactly, a binomial experiment. The reason is that, for instance, after selecting one of the female fish, the remaining numbers in the tank are 274 females and 334 males. Hence the probability of selecting a female the next time would be $\frac{274}{608}$, which is slightly smaller than $\frac{275}{609}$. Similarly, the probability of selecting a male would be $\frac{334}{608}$ which is slightly larger than $\frac{334}{609}$. There is another family of discrete random variables known as the *hypergeometric* family that models selections done *without replacement*. The hypergeometric p.m.f. would yield the value

$$P(Y = 6) = \frac{\binom{275}{6} \cdot \binom{334}{4}}{\binom{609}{10}} \doteq .1613.$$

This is close to, but not exactly the same as, the probability .1611 we obtained from the binomial experiment for the probability of selecting 6 females out of the 10 fish.

To write the hypergeometric p.m.f. in general, we will use the language from Example 10.7. If we draw n balls *without replacement* from an urn containing N balls, of which n_1 are red and n_2 are black, then the probability that we obtain exactly k red balls is given by

$$\boxed{P(k \text{ red balls}) = \frac{\binom{n_1}{k} \cdot \binom{n_2}{n-k}}{\binom{N}{n}}.} \tag{10.7}$$

The numerator counts the total number of ways of picking k of the red balls and $n-k$ of the white balls. The denominator is the total number of ways to pick n balls from the N in the urn. Hence the formula follows from the basic ratio methods of finding theoretical probabilities from Chapter 2. It can be shown that this is close to the binomial probability with $p = \frac{n_1}{N}$ and $q = \frac{n_2}{N}$ when N, n_1, n_2 are all large (as was true for instance in the example discussed above). However, when N is small, this can differ significantly from the binomial probability.

Geometric Distributions and Probability Mass Functions. Now suppose we are in a similar situation to the binomial experiments described above, but with two important differences. Namely,

(1) we do Bernoulli trials, all with the same success probability p, but the number of trials is no longer fixed,
(2) the trials are *independent* in the sense discussed above, and
(3) the random variable of interest is $Y =$ the *number of the trial on which the first success is observed*.

In this situation, we have $Y = k$ exactly when trials $1, 2, \ldots, k-1$ produced failure outcomes, but trial number k produced the first success. Hence by the independence of the individual trials, we see the probability is a product of $k-1$ factors of q and one final factor of p:

$$\boxed{P(Y = k) = q \cdot q \cdots q \cdot p = q^{k-1}p \text{ for all } k \geq 1.} \tag{10.8}$$

Random variables having a probability mass function of the form given in (10.8) are said to have a *geometric* distribution. Unlike binomial random variables, geometric random variables have $P(Y = k)$ nonzero for all $k \geq 1$. Even for very large k, there is a small but strictly positive chance that the first success occurs on trial k.

The reason for the name "geometric random variables" becomes somewhat clearer if we think of summing the probabilities in (10.8) for $k \geq 1$:

$$\sum_{k=1}^{\infty} q^{k-1}p = p + pq + pq^2 + pq^3 + \cdots . \tag{10.9}$$

This is an (infinite) geometric series, that is, a series of the form

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + ar^3 + \cdots .$$

When $-1 < r < 1$ (or equivalently $|r| < 1$), this sum converges to (that is, "adds up to")[9] the sum given in this formula:

$$\boxed{\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}.} \tag{10.10}$$

When $|r| \geq 1$, the series does not add up to a finite sum. Hence the series in (10.9), where $a = p$ and $r = q$, adds up to the value

$$\frac{p}{1-q} = \frac{p}{p} = 1,$$

and we obtain the sum 1 as we expect. We have covered *all of the possible outcomes* by including the terms with $k = 1, 2, 3, \ldots$, so we should obtain 1 by summing all of those probabilities.

[9]See the (optional) Exercise 15 below for the derivation of this fact.

EXAMPLE 10.10. Suppose that a prey organism lives in a habitat with predators, and each prey individual has a 10% chance of being killed by a predator or dying from other causes each month. We will assume the prey individuals are solitary, so that the fate of one of the animals has no effect on the fates of the others.[10] We will also assume that whether an animal survives the predation in one month is independent of whether they survive in other months. Then we can model their survival probabilities using geometric random variables. Here the "success" probability as we have described it would actually correspond to the "bad" outcome of being eaten by a predator or dying from other causes, so our previous caveat about interpreting the words "success" and "failure" needs to be taken seriously.[11]

(a) What is the probability that a prey individual survives 4 months before being eaten in the month 5? This is $P(Y = 5) = (.9)^4 \cdot (.1) = .06561$, or less than a 7% chance.

(b) What is the probability that a prey individual survives *at least four months*? This means that we want

$$P(Y \geq 5) = \sum_{k=5}^{\infty} P(Y = k) = (.9)^4 \cdot (.1) + (.9)^5 \cdot (.1) + (.9)^6 \cdot (.1) + \cdots .$$

We see that this is another infinite geometric series with ratio $r = .1$, but now the first term is $(.9)^4 \cdot (.1)$. Hence, by the formula from Exercise 10.10, the sum is

$$= \frac{(.9)^4 \cdot (.1)}{1 - .9} = .6561,$$

or roughly a 66% chance. △

Poisson Distributions and Probability Mass Functions. Binomial random variables count the number of "successes" in a fixed number of independent Bernoulli trials and geometric random variables deal with the number of the trial on which the first success is observed. The final type of discrete random variable we will consider, the so-called *Poisson random variables*, are used to model *the probabilities of different numbers of occurrences of a particular event within a fixed region in time or space.* Here is a typical example of the type of problem that leads to Poisson random variables.

EXAMPLE 10.11. Suppose that a forester knows that maple seedlings are randomly dispersed through a certain area of the forest, with an average density of 2 seedlings per square meter. If the forester marks out five one meter square plots in the forest, what is the probability that *none* of the plots contains a maple seedling? Similarly, what is the probability that the five plots together contain exactly two seedlings? △

A discrete random variable Y is said to have a *Poisson distribution* if its p.m.f. has the form

$$\boxed{P(Y = k) = \frac{\lambda^k}{k!} e^{-\lambda},} \tag{10.11}$$

[10] This is actually rarely true in real life, of course – prey animals tend to live in larger groups for "safety in numbers."

[11] Or alternatively, we might look at things from the point of view of the predators!

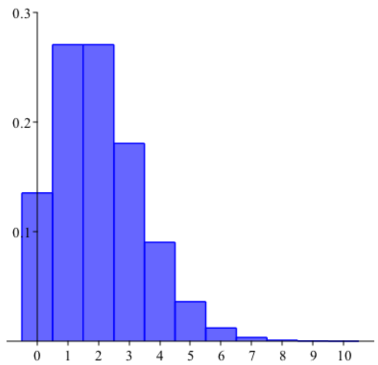


FIGURE 10.2. Poisson probabilities, $\lambda = 2$.

where $\lambda > 0$ is a constant representing the *average value* of Y (in a sense to be made precise in the next section), and $e^{-\lambda}$ is a value of the natural exponential function introduced in Chapter 5, for which the base is the number $e \doteq 2.71828\cdots$. This formula for the Poisson p.m.f. can be derived from the binomial p.m.f. by a sort of limiting process (subdividing the interval in time or space into subintervals small enough that either 1 or 0 events occur). We will not discuss the details of this derivation in the main text, though, since they rely on some results typically only established in calculus courses. See the (optional) Exercise 16 below.

EXAMPLE 10.12. We continue with the question introduced in Example 10.11. Here it is the number of seedlings in a one meter square area that would be the random variable Y with a Poisson distribution. The parameter $\lambda = 2$, since we are given that the average density of maple seedlings is 2 seedlings per square meter. The probability mass function *for the number of seedllings in a one square meter area* is then

$$P(Y = k) = \frac{2^k}{k!}e^{-2}$$

for all $k \geq 0$. The Poisson random variables are like geometric random variables in that there is a strictly positive probability that Y has each possible value $k \geq 0$, although those probabilities drop off quickly as k increases. See the Poisson probability histogram plotted in Figure 10.2. The probability that a one meter square plot contains no seedlings is

$$P(Y = 0) = \frac{2^0}{0!}e^{-2} \doteq .1353.$$

Hence the probability that none of the five plots contains a seedling is (assuming independence) $(.1353)^5 \doteq 4.54 \times 10^{-5}$ (in other words, virtually no chance of that happening!)

The second part of the question was to determine what the probability was that the 5 one meter square plots contain exactly two seedlings between them. Note that this could happen in several ways. One of the plots could contain two seedlings, while the other four had none. Since any one of the five plots could be the one containing the two seedlings, the probability of that happening would be

$$5 \cdot P(Y=2) \cdot P(Y=0)^4 = 5 \cdot \frac{2^2}{2!} e^{-2} \cdot \left(\frac{2^0}{0!} e^{-2} \right)^4 \doteq .000454.$$

Next, there are $\binom{5}{2} = 10$ ways for 2 plots to have one seedling each and the others to have none. This gives a contribution of

$$10 \cdot P(Y=1)^2 \cdot P(Y=0)^3 = 10 \cdot \left(\frac{2^1}{1!} e^{-2} \right)^2 \cdot \left(\frac{2^0}{0!} e^{-2} \right)^3 \doteq .001816.$$

Adding, we obtain approximately

$$\text{Probability of two seedlings} = .000454 + .001816 = .002270 \tag{10.12}$$

(about a 0.2% chance – we would not expect to see this often!).

Before leaving this example, we want to make another observation about what is happening here. Note that if the 5 one meter square plots were *contiguous*, then we might immediately notice that the average number of seedlings in the 5 square meters would be $5 \cdot 2 = 10$. But that is true in any case for *any* region consisting of 5 square meters of the forest. Moreover, the Poisson formula with $\lambda = 10$ would give

$$P(Y=2) = \frac{10^2}{2!} e^{-10} \doteq .002270,$$

which is exactly the same number as in (10.12) above before and after rounding. In fact, *the Poisson p.m.f. formula always applies in exactly the same way, whether the region in space or time is contiguous or not.* We only need to know the appropriate value of the parameter λ for the whole region in order to apply it. △

10.4. Expected Value and Variance

In the previous chapter we discussed computing the mean and the standard deviation of a collection of numerical data. In this section we will turn to the corresponding concepts for discrete random variables described by probability mass functions (p.m.f.'s) as in the last section. The concept corresponding to the mean is usually called the *expected value* of a random variable, and the concept corresponding to the standard deviation goes by that name in this context too. But in fact we will begin by showing how to compute the *variance* of a discrete random variable and the standard deviation will be obtained by taking the square root of the variance.

Expected Values. The expected value of a discrete random variable is computed as a *weighted average* of the values. The weights are the associated probabilities or p.m.f. values. Hence we can write a general formula for computing the expected value in the following form, when it is true that the values are a subset of the integers ≥ 0 (i.e. not just some set in one-to-one correspondence with the natural numbers). The notation we will use for this is $E(Y)$:

$$\boxed{E(Y) = \textstyle\sum_{k=0}^{\infty} k \cdot P(Y=k).} \tag{10.13}$$

If $P(Y=k)$ is zero except for a finite collection of k values, this is an ordinary finite sum. In other cases, it will be an infinite series and some care is needed to determine whether the sum converges to a finite result. To see what this really means, we will work out a few examples.

EXAMPLE 10.13. As is true for many of the concepts in probability theory, the idea of the expected value goes back to the analysis of outcomes in the *games of chance* that people have played in most cultures throughout history.[12] For instance, suppose you play a simple lottery game where you buy a \$1 ticket along with 999 other people (1000 total tickets sold). There are five \$10 prizes and one \$100 grand prize. We could ask: What is our expected return from the \$1 purchase?

Assume the grand prize winner is drawn first and the five other winners are chosen at random (without replacement) from the remaining 999 people. By (10.7), if you do not win the grand prize, you still have probability of

$$\frac{\binom{1}{1} \cdot \binom{998}{4}}{\binom{999}{5}} = \frac{5}{999}$$

of winning one of the other prizes.

You have a random variable $R =$ your return in dollars from the lottery with

$$\begin{aligned} P(R = 100) &= \frac{1}{1000} = .001, \\ P(R = 10) &= \frac{5}{999} \doteq .005, \text{ and} \\ P(R = 0) &= 1 - \frac{1}{1000} - \frac{5}{999} \doteq .994. \end{aligned}$$

According to (10.13), your expected return is

$$E(R) \doteq 0 \cdot .994 + 10 \cdot .005 + 100 \cdot .001 \doteq .15$$

or about \$0.15.[13] $\triangle$

EXAMPLE 10.14. Consider the binomial random variable X with $n = 5$ and $p = \frac{1}{3}$. The values of the p.m.f. are given in Table 1. According to the formula

TABLE 1. A binomial p.m.f. with $n = 5$ and $p = \frac{1}{3}$

k	0	1	2	3	4	5
$P(Y = k)$	$\frac{32}{243}$	$\frac{80}{243}$	$\frac{80}{243}$	$\frac{40}{243}$	$\frac{10}{243}$	$\frac{1}{243}$

[12] The mathematical treatment of probabilities in games of chance goes back to works of Blaise Pascal, René Descartes and others. Jacob Bernoulli's *Ars Conjectandi*, first published in 1713, put the theory in something close to its modern form. Bernoulli trials are named after him.

[13] The organizations that run lotteries (often state governments, charitable organizations such as churches, etc.) are not in that business out of the kindness of their hearts or for the entertainment of those who gamble on winning; they are usually doing it to *make money*.

(10.13) above, we want to compute

$$\begin{aligned}E(Y) &= 0 \cdot P(Y=0) + 1 \cdot P(Y=1) + 2 \cdot P(Y=2) + 3 \cdot P(Y=3) \\ &\quad + 4 \cdot P(Y=4) + 5 \cdot P(Y=5) \\ &= 0 \cdot \frac{32}{243} + 1 \cdot \frac{80}{243} + 2 \cdot \frac{80}{243} + 3 \cdot \frac{40}{243} \\ &\quad + 4 \cdot \frac{10}{243} + 5 \cdot \frac{1}{243} \\ &= \frac{80 + 160 + 120 + 40 + 5}{243} \\ &= \frac{405}{243} \\ &= \frac{5}{3},\end{aligned}$$

since $405 = 5 \cdot 81$ and $243 = 3 \cdot 81$.

In hindsight, this result should look almost “obvious” – we are doing five identical Bernoulli trials with a success probability of $p = \frac{1}{3}$ on each trial. So the expected value of the number of successes could hardly be anything other than $np = \frac{5}{3}$(!) $\triangle$

The same pattern holds in general.

THEOREM 10.15. *Let Y be a binomial random variable based on n trials, with success probability p. Then $E(Y) = np$.*

The proof consists of some clever manipulations of the formula

$$E(Y) = \sum_{k=0}^{n} \left(k \cdot \binom{n}{k} p^k q^{n-k} \right),$$

but we will not go into the details here because the intuition gained through the interpretation given at the end of Example 10.14 is more important.

EXAMPLE 10.16. To find the expected value for a geometric random variable using (10.13), we need to sum

$$E(Y) = \sum_{k=1}^{\infty} k \cdot q^{k-1} p = p(1 + 2q + 3q^2 + 4q^3 + \cdots).$$

We can identify the sum in the parentheses if we think of breaking it up into several geometric series, summed vertically then horizontally as in the following diagram:

$$\begin{array}{ccccc} 1 & +q & +q^2 & +q^3 & +q^4 \\ & +q & +q^2 & +q^3 & +q^4 \\ & & +q^2 & +q^3 & +q^4 \\ & & & +q^3 & +q^4 \\ & & & & \ddots \end{array}$$

If we reorder the sum (this is permissible in this case provided $0 < q < 1$) by adding across each row, then adding the row sums, we can apply (10.10) to obtain

$$\frac{1}{1-q} + \frac{q}{1-q} + \frac{q^2}{1-q} + \cdots = \frac{1 + q + q^2 + \cdots}{1-q} = \frac{1}{(1-q)^2},$$

using (10.10) once more. Hence the formula for the expected value of the geometric random variable Y is

$$E(Y) = \frac{p}{(1-q)^2} = \frac{1}{p}$$

since $1 - q = p$. Again, this should look "almost obvious" with the benefit of hindsight. If we have a success probability of p, then it should take $\frac{1}{p}$ trials on average before we get a success. For instance, with $p = \frac{1}{5}$, then we would expect roughly one success every 5 trials, so on average, we might expect it to take 5 trials to get the first success. $\triangle$

THEOREM 10.17. *Let Y be a geometric random variable with success probability p. Then $E(Y) = \frac{1}{p}$.*

Finally we consider the Poisson random variables introduced in section 10.2. For a Poisson variable with parameter λ, from (10.13) we obtain

$$E(Y) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda}.$$

By canceling the k in the numerator with one of the factors in the $k!$ in the denominator, this can be rewritten as

$$= \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}.$$

If we rewrite the sum using a new index $\ell = k - 1$, then we obtain

$$= \lambda \cdot \left(\sum_{\ell=0}^{\infty} \frac{\lambda^{\ell}}{\ell!} e^{-\lambda} \right).$$

Now the sum in the parentheses on the right is just the sum of the values of the p.m.f. of the Poisson random variable. *Hence it must equal* 1.[14] As a result, we see:

THEOREM 10.18. *Let Y be a Poisson random variable with parameter λ. Then $E(Y) = \lambda$.*

This agrees with the way we described the meaning of the parameter λ when we introduced the Poisson random variables in the last section.

Variances. The variance of a random variable is perhaps best understood as the expected value of a certain *function of the random variable*. Namely we consider the weighted average of the *square deviation from the expected value* (as in the discussion of the formula for the standard deviation from the last chapter), again weighted by the p.m.f. This defines the variance of a discrete random variable with values in the non-negative integers, denoted $V(Y)$:

$$\boxed{V(Y) = E((Y - E(Y))^2) = \textstyle\sum_{k=0}^{\infty} (k - E(Y))^2 P(Y = k).} \tag{10.14}$$

[14]Technical note: This fact can also be seen from the fact that the sum $\sum_{\ell=0}^{\infty} \frac{\lambda^{\ell}}{\ell!}$ is the Taylor series for the exponential e^{λ}, so we have $e^{\lambda} \cdot e^{-\lambda} = 1$.

EXAMPLE 10.19. Consider the binomial random variable with p.m.f. given in Table 1. According to (10.14), the variance will be

$$\begin{aligned}
V(Y) = &\left(0-\frac{5}{3}\right)^2 \cdot P(Y=0) + \left(1-\frac{5}{3}\right)^2 \cdot P(Y=1) + \left(2-\frac{5}{3}\right)^2 \cdot P(Y=2) \\
&+ \left(3-\frac{5}{3}\right)^2 \cdot P(Y=3) + \left(4-\frac{5}{3}\right)^2 \cdot P(Y=4) + \left(5-\frac{5}{3}\right)^2 \cdot P(Y=5) \\
= &\frac{25}{9}\cdot\frac{32}{243} + \frac{4}{9}\cdot\frac{80}{243} + \frac{1}{9}\cdot\frac{80}{243} + \frac{16}{9}\cdot\frac{40}{243} + \frac{49}{9}\cdot\frac{10}{243} + \cdot\frac{100}{9}\cdot\frac{1}{243} \\
= &\frac{800+320+80+640+490+100}{9\cdot 243} \\
= &\frac{2430}{9\cdot 243} \\
= &\frac{10}{9}.
\end{aligned}$$

This value is perhaps less “obvious” than what we saw before for the expected value of Y. Nevertheless, note that it is equal to $5\cdot\frac{1}{3}\cdot\frac{2}{3} = npq$. We could also define a standard deviation $\sigma(Y)$ by taking the square root:

$$\sigma(Y) = \sqrt{\frac{10}{9}} = \frac{\sqrt{10}}{3}.$$

We will discuss interpretations of the standard deviation for a random variable later. △

The pattern seen in Example 10.19 is also true in general for binomial random variables.

THEOREM 10.20. *Let Y be a binomial random variable based on n trials, with success probability p. Then $V(Y) = npq$.*

This and the corresponding formulas for geometric and Poisson random variables can be shown in general by more involved summation manipulations along the lines of the ones we have seen so far. The final results of a whole lot of algebra are given in the following statements.

THEOREM 10.21. *Let Y be a geometric random variable with success probability p. Then $V(Y) = \frac{q}{p^2}$.*

THEOREM 10.22. *Let Y be a Poisson random variable with parameter λ. Then $V(Y) = \lambda$.*

10.5. Continuous Distributions and Probability Density Functions

We now turn to considering what are known as *continuous* random variables and their probability distributions. This is a topic where a complete discussion requires techniques from integral calculus, so we will only sketch the main ideas and indicate some representative examples.

To begin, continuous random variables can take any value in an interval on the real number line, an uncountably infinite set of values. To describe them, we use

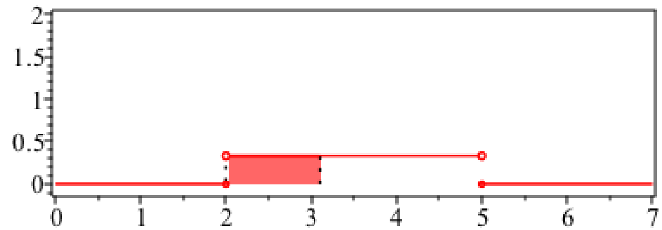


FIGURE 10.3. p.d.f. of a uniform random variable on $2 \leq t \leq 5$.

what are known as *probability density functions* or p.d.f.s for short. The power law distributions considered in Chapter 6 can be thought of as examples, where the p.d.f. is a power function and nonzero for all positive real numbers. For reasons that should become clearer soon, we will use the variable name t for the input to the p.d.f. of a continuous random variable. The idea is to consider something like the probability histograms we saw in Figures 10.1 and 10.2, but where the collection of boxes centered on the integer t-values is replaced by a general region between a graph $y = f(t)$ and the t-axis. The function $f(t)$ giving the graph is the p.d.f. and it must satisfy the following properties:

(1) $f(t) \geq 0$ for all real t, and
(2) The total area between the graph $y = f(t)$ and the t-axis is equal to 1.[15]

For a continuous random variable X with p.d.f. equal to $f(t)$ we define:[16]

$$\boxed{P(a \leq X \leq b) = \text{total area between } t\text{−axis and } y = f(t), a \leq t \leq b.} \tag{10.15}$$

Because of conditions 1 and 2 above, this implies

$$0 \leq P(a \leq X \leq b) \leq 1$$

for all real numbers a, b and we can interpret the numbers we get as in the discrete case.

The p.d.f. also makes it possible to compute what is known as the *cumulative distribution function*, or c.d.f. of a continuous random variable.[17] The c.d.f. is defined as

$$\boxed{F(x) = P(X \leq x) = \text{total area between } t\text{−axis and } y = f(t), t \leq x.} \tag{10.16}$$

We use the name t for the variable in the p.d.f. here to distinguish it from the x which is the input to the c.d.f. $F(x)$.

Here is a simple, but representative example.

---

[15]Technical note: This is a first place where integral calculus comes into play. The precise way to state this is that we require $\int_{-\infty}^{\infty} f(t)\, dt = 1$.

[16]Technical note: Again, the precise meaning uses calculus: $P(a \leq X \leq b) = \int_a^b f(t)\, dt$.

[17]Technical note: Continuity of a random variable is usually defined in more advanced treatments of probability theory by saying that the c.d.f. of the random variable is a continuous function. The c.d.f. of a discrete random variable with values in the integers can also defined, but it will have a jump discontinuity at each integer k where $P(X = k) \neq 0$. There are also random variables that can be seen as "mixtures" of the continuous and discrete types. We will not consider them.

EXAMPLE 10.23. Continuous random variables with a p.d.f. of the form

$$f(t) = \begin{cases} \frac{1}{d-c}, & c < t < d \\ 0, & \text{otherwise} \end{cases} \tag{10.17}$$

for some $c < d$ are known as *uniform* random variables.

For any uniform random variable, the probability that X lies in an interval (a, b) contained in (c, d) is simply equal to the area of the *rectangle* under the graph $y = f(t)$ for $a < t < b$. The corresponding area is the ratio

$$P(a < X < b) = (b - a) \cdot \frac{1}{c - d} = \frac{b - a}{c - d}.$$

If (a, b) does not lie completely within (c, d), then we would take the area under $y = \frac{1}{d-c}$ for t in the part of (a, b) that does lie within (c, d).

For instance, if $c = 2$ and $d = 5$, then the graph of the p.d.f. is shown in Figure 10.3. The open circles at $y = \frac{1}{3}$ and the solid circles at $y = 0$ indicate that the values $f(2)$ and $f(5)$ are actually equal to 0 according to this definition. For a random variable Y with that p.d.f., we would have $P(3 < Y < 4) = (4-3)\cdot\frac{1}{5-2} = \frac{1}{3}$. Similarly, $P(Y > 4.5) = \frac{1}{6}$, since the part of the interval of $t > \frac{9}{2}$ within $(2, 5)$ has length $\frac{1}{2}$, so we are looking at the area of a rectangle of width $\frac{1}{2}$ and height $\frac{1}{3}$.

The shaded red rectangle here is the region under the graph of the p.d.f. between $t = 2$ and $t = 3.1$, which is the same as the region for $t \le 3.1 = \frac{31}{10}$. We see from this that

$$P\left(2 \le x \le \frac{31}{10}\right) = \frac{1}{3} \cdot \left(\frac{31}{10} - 2\right) = \frac{11}{30} \doteq .3667.$$

Similarly, since $f(t) = 0$ for all $t \le 2$, we have

$$F(3.1) = P(X \le 3.1) = \frac{11}{30} \doteq .3667.$$

If we imagine x beginning to the left of $t = 2$, sweeping to the right past $t = 2$, through the interval (including the value $t = 3.1$ shown above in Figure 10.3, then through $t = 5$ and beyond, we can see that the value of the c.d.f. $F(x)$ will be equal to 0 for all $x \le 2$, then increase steadily and linearly (we're adding area at a constant rate if x increases at a constant rate) from 0 up to 1 on the interval $2 \le x \le 5$, then remain at 1 for all $x \ge 5$:

$$F(x) = \begin{cases} 0, & x \le 2 \\ \frac{1}{3}(x - 2), & 2 < x < 5 \\ 1, & x \ge 5. \end{cases}$$

The c.d.f. is shown in Figure 10.4. Note that this graph is continuous even though the p.d.f. had jump discontinuities at $t = 2$ and $t = 5$. As is true in this example, c.d.f.s of continuous random variables are always "non-decreasing" – that is $x_1 \le x_2$ always implies $F(x_1) \le F(x_2)$. $\triangle$

Verifying the details claimed in the following example requires some calculus, but we will present it in any case because it illustrates the connection between p.d.f.s and c.d.f.s even more clearly.

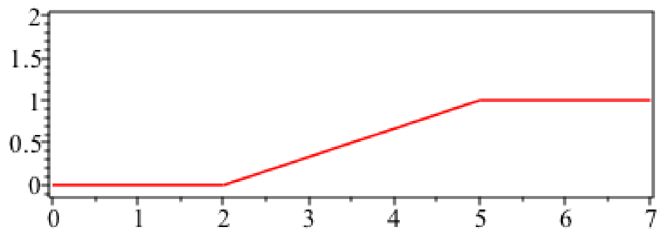


FIGURE 10.4. c.d.f. of the uniform random variable from Example 10.23.

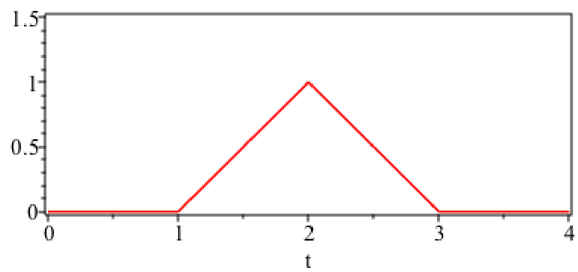


FIGURE 10.5. p.d.f. from Example 10.24.

EXAMPLE 10.24. The graph of the "tent function"

$$f(t) = \begin{cases} 0, & t \leq 1 \\ t-1, & 1 < t \leq 2 \\ -t+3, & 2 < t \leq 3 \\ 0, & t > 3 \end{cases}$$

is shown in Figure 10.5. This is a valid p.d.f. because $f(x) \geq 0$ for all real x, and the total area between the graph and the x-axis is 1 since it is made up of two triangles with base 1 and height 1.

The corresponding c.d.f. is given by

$$F(x) = \begin{cases} 0, & x \leq 1 \\ \dfrac{x^2}{2} - x + \dfrac{1}{2}, & 1 < x \leq 2 \\ \dfrac{-x^2}{2} + 3x - \dfrac{7}{2}, & 2 < x \leq 3 \\ 1, x > 3. \end{cases}$$

This function is shown in Figure 10.6.

Readers who know some calculus might recognize that in both of the cases we have worked out, the c.d.f. has the property that *its derivative is the corresponding*

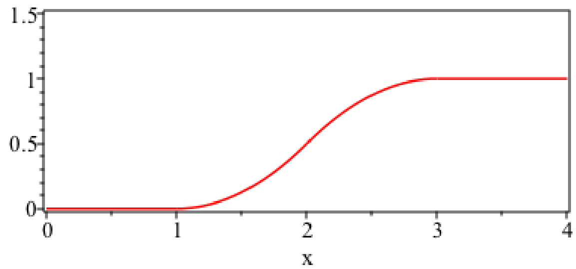


FIGURE 10.6. c.d.f. from Example 10.24.

p.d.f.[18] Even if you have not seen these ideas from calculus before, note how the graph of the c.d.f. is "smoother" than the graph of the p.d.f. in both cases. △

The formulas for computing expected values and variances from Section 10.3 generalize to the case of continuous random variables as well. However, this is one place where tools from calculus are actually very necessary for computations in almost all cases.[19] For instance, for the uniformly distributed random variables, we have:

THEOREM 10.25. *Let Y be a random variable with a uniform distribution on the interval (c, d). Then*

$$E(Y) = \frac{c+d}{2} \qquad \text{and} \qquad V(Y) = \frac{(d-c)^2}{12}.$$

In this case, the expected value is just the midpoint of the interval (c, d) where the density is nonzero. The variance is proportional to the square of the length of the interval and hence the standard deviation is proportional to the length of the interval, corresponding to the intuitively clear idea that the larger $d-c$ is, the more variability there will be in values of the random variable.

To conclude this section, we will state without proof an important fact about collections of random variables $Y_1, \ldots, Y_n$ satisfying the following two properties:

- We will assume that the Y_i all have *the same* distribution (i.e. the same p.m.f. in the discrete case or the same p.d.f. in the continuous case),

[18] Technical note: This is a consequence of the first part of the Fundamental Theorem of Calculus:

$$F'(x) = \frac{d}{dx}\int_{-\infty}^{x} f(t)\ dt = f(x).$$

[19] Technical note: The expected value of the continuous random variable X with p.d.f. $f(x)$ is computed by

$$E(Y) = \int_{-\infty}^{\infty} x \cdot f(x)\ dx$$

and the variance is

$$V(Y) = E((Y - E(Y))^2) = \int_{-\infty}^{\infty} (x - E(Y))^2 f(x)\ dx.$$

- We will assume that the Y_i are *independent* in the sense that

$$P(a_i < Y_i < b_i \text{ and } a_j < Y_j < b_j) = P(a_i < Y_i < b_i) \cdot P(a_j < Y_j < b_j)$$

for all pairs $i \neq j$ and all intervals (a_i, b_i) and (a_j, b_j).[20]

These conditions are often abbreviated by saying the Y_i are i.i.d. ("independent and identically distributed") random variables.

THEOREM 10.26. *Let $Y_1, \ldots, Y_n$ be an i.i.d. collection of random variables with $E(Y_i) = \mu$ and $V(Y_i) = \sigma^2$ for some μ and σ. Let $\overline{Y} = \frac{1}{n}(Y_1 + \cdots + Y_n)$. Then $\overline{Y}$ is a new random variable with*

$$E(\overline{Y}) = \mu \qquad \text{and} \qquad V(\overline{Y}) = \frac{\sigma^2}{n}.$$

The statement about $V(\overline{Y})$ is the ultimate explanation for what you saw in question G in the Chapter Project from Chapter 9. Recall that the SDs of the sample means seemed to decrease as the sample size increased. In fact, the exact pattern is that the means of samples of size n will have $SD = \frac{\sigma}{\sqrt{n}}$.

10.6. Normal Distributions

The most important continuous probability distributions in traditional methods for analyzing scientific data and making statistical inferences are certain the *normal distributions*. A continuous random variable Y is said to have a normal distribution if its p.d.f. has the form

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t-\mu)^2/(2\sigma^2)}. \tag{10.18}$$

This formula includes two parameters μ and σ and it can be shown, via tools from calculus, that $\mu = E(Y)$ and $V(Y) = \sigma^2$, so changing μ changes the expected value, and changing σ changes the variance and hence the standard deviation. The constant multiplier $1/\sqrt{2\pi\sigma^2}$ is included to make the total area between the graph $y = f(t)$ and the t-axis equal to 1.

Among the normal p.d.f.s, a special role is played by the function of the form from (10.18) with $\mu = 0$ and $\sigma = 1$:

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}, \tag{10.19}$$

called the *standard normal p.d.f.*. This function is plotted for $-4 \leq t \leq 4$ in Figure 10.7. This is the standard "bell-shaped" curve that defines normally-distributed random variables.

Using numerical techniques, we can compute the area of the region between this graph and the t-axis for $-4 \leq t \leq 4$, and the result is approximately .99994, very close to 1, but a bit less. The "missing area" lies in the *tails of the distribution*: $t > 4$ and $t < -4$.

Changing μ while leaving σ unchanged merely shifts the whole p.d.f. plot to the right or left. Changing σ without changing μ changes the height and the width of the *peak*, or spike, of the graph at $x = \mu$. For instance, a large value of σ will

[20]Technical note: Independence is usually defined using the idea of the *joint p.d.f.* of the collection of random variables and this statement about probabilities is a consequence.

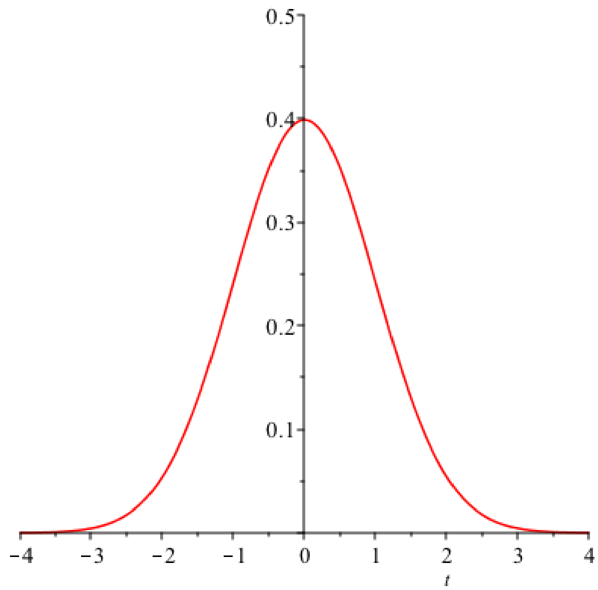


FIGURE 10.7. The standard normal p.d.f.

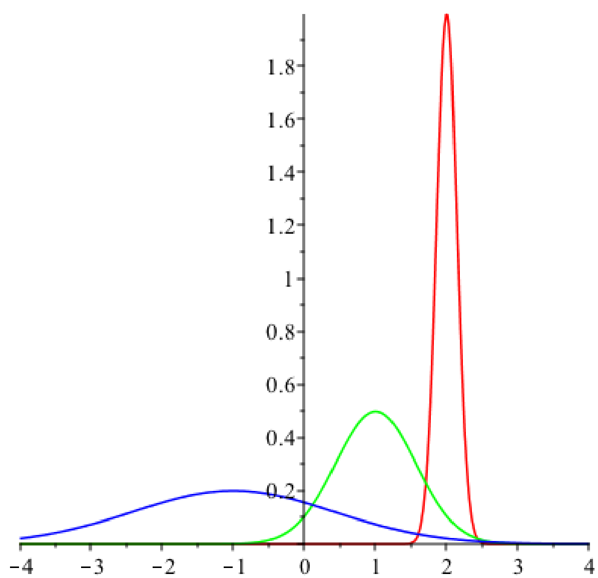


FIGURE 10.8. Several normal p.d.f.s plotted together

give a short, widely-spread peak, while a small value of σ will yield a tall, narrow peak. Several examples plotted together on the same set of axes in Figure 10.8. The red graph is the normal p.d.f. with $\mu = 2$, $\sigma = .2$; the green graph has $\mu = 1$ and $\sigma = .8$, while the blue graph has $\mu = -1$ and $\sigma = 2$.

We will introduce several other continuous distributions related to the normals, including χ^2 and t distributions in future chapters.

10.7. Z-scores and Computing Normal Probabilities

The reason that the standard normal from (10.19) and Figure 10.7 is so important here is the following fact, shown in more advanced probability and mathematical statistics courses.

Theorem 10.27. *If Y has a normal distribution with expected value μ and standard deviation σ, then*

$$Z = \frac{Y - \mu}{\sigma}$$

has a standard normal distribution.

This means that any computation of areas under any normal p.d.f. graph can be translated into an equivalent problem using the standard normal p.d.f.. The idea is the following. If Y is a normal random variable with expected value μ and standard deviation σ, then the inequalities $a \le Y \le b$ can be converted into an *equivalent pair of inequalities* by subtracting μ and dividing through by the positive number σ:

$$\frac{a-\mu}{\sigma} \le \frac{Y-\mu}{\sigma} \le \frac{b-\mu}{\sigma}.$$

By the theorem the middle term here is a standard normal, and we have reduced the problem to finding an area between the graph of the standard normal p.d.f. graph and the horizontal axis. Moreover,

$$\boxed{P(a \le Y \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)}, \tag{10.20}$$

because the two sets of inequalities are equivalent.

If y is a value of the normally-distributed random variable Y (with expected value μ and standard deviation σ), the number

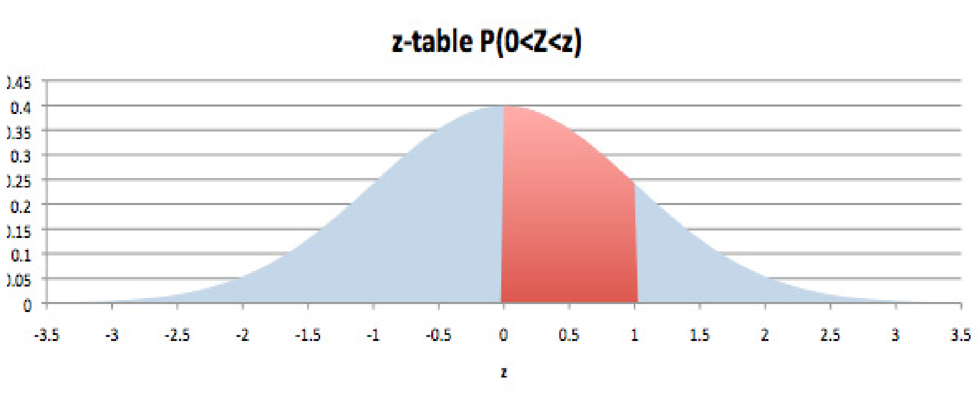
$$z = \frac{y-\mu}{\sigma}$$

is often called the *Z-score corresponding to y*. So (10.20) says that to find the probability that a normal random variable takes a value in a given interval, we simply compute the Z-scores corresponding to the endpoints of the interval and find the area under the standard normal for the interval with endpoints at the Z-scores.

Traditionally this is done by way of a *table of standard normal curve areas*[21], although use of mathematical and statistical software for this purpose is also becoming more and more common.[22] One such table is given in Figure 10.9 on the next page.

[21]From: `http://statistics.wdfiles.com/local--files/ch7/normDistTable.pdf`. License: CC BY-SA: Attribution-ShareAlike, downloaded October 10, 2017.

[22]Technical note: the standard normal p.d.f. does not have an elementary antiderivative, so symbolic computation with the Fundamental Theorem of Calculus is insufficient for finding the corresponding integrals. The computation of these integrals is done by approximate numerical



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990
3.1	.4990	.4991	.4991	.4991	.4992	.4992	.4992	.4992	.4993	.4993
3.2	.4993	.4993	.4994	.4994	.4994	.4994	.4994	.4995	.4995	.4995
3.3	.4995	.4995	.4995	.4996	.4996	.4996	.4996	.4996	.4996	.4997
3.4	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4997	.4998
3.5	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998	.4998
	.00	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**

FIGURE 10.9. Table of values of $P(0 < Z < z)$

To learn how to use the table, note the heading showing another plot of the standard normal p.d.f. with the shaded region starting at $t = 0$ and extending to the right until $t = z$. The ones digit and the tenths digit of the number $z > 0$ are given along the left column and the hundredths digit of z is given across the top and again at the bottom of the table.

For example, to find the area representing $P(0 < Z < 1.21)$, we would look in the row labeled 1.2 and the column labeled .01 to find

$$P(0 < Z < 1.21) \doteq .3869.$$

Similarly the area representing $P(0 < Z < 0.47)$ is found in the row labeled 0.4 and the column labeled .07:

$$P(0 < Z < 0.47) \doteq .1808.$$

Note that by the time z reaches $z = 3.59$ the area has "settled down" to a number very close to .5000, so the table does not continue to higher values of z.

By adding or subtracting areas, and/or making use of the *symmetry* of the standard normal p.d.f. ($f(-t) = f(t)$) we can compute many other types of areas under the standard normal curve besides those of the form $P(0 < Z < z)$. The following discussion covers (almost) all of the cases that arise in practice:

(1) By the symmetry $P(Z > 0) = P(Z < 0)$ and each of these $= .5000$, since the total area must equal 1.0.

(2) Hence if $z > 0$ the *upper tail area* $P(Z > z)$ equals the difference between $P(Z > 0) = .5000$ and the area between 0 and z:

$$P(Z > z) = P(Z > 0) - P(0 < Z < z) = .5000 - P(0 < Z < z).$$

For example, $P(Z > .89) = .5000 - P(0 < Z < .89) = .5000 - .3133 = .1867$.

(3) The area of a region extending from a negative number z to 0 can be obtained since $P(-z < Z < 0) = P(0 < Z < z)$. For example,

$$P(-2.30 < Z < 0) = P(0 < Z < +2.30) \doteq .4893.$$

(4) Combining (2) and (3) we see that when $z < 0$ the *lower tail area* $P(Z < z) = P(Z > -z) = .5000 - P(0 < Z < -z)$ (note $-z > 0$ if $z < 0$ as we are assuming here). For instance $P(Z < -.63) = P(Z > .63) = .5000 - .2357 = .2643$.

(5) For areas extending from one positive number $z_1 > 0$ to another $z_2 > z_1$, such as $P(z_1 < Z < z_2)$, we can obtain the desired area by subtraction:

$$P(z_1 < Z < z_2) = P(0 < Z < z_2) - P(0 < Z < z_1).$$

For example,

$$\begin{aligned} P(0.48 < Z < 1.64) &= P(0 < Z < 1.64) - P(0 < Z < 0.48) \\ &= .4495 - .1844 \\ &= .2651. \end{aligned}$$

Applying the same idea and symmetry, an area $P(z_1 < Z < z_2)$ with $z_1 < z_2 < 0$ can also be obtained easily.

---

integration and the table in the text and similar tables available in many places are all generated in that way.

(6) Finally, for intervals containing both positive and negative numbers, we can split the interval at zero and sum intervals of the types given in the previous parts to find the desired areas. For example,

$$\begin{aligned} P(-.37 < Z < 1.28) &= P(-.37 < Z < 0) + P(0 < Z < 1.28) \\ &= P(0 < Z < .37) + P(0 < Z < 1.28) \end{aligned}$$

(by the symmetry). From the table we find this equals $.1443 + .3997 = .5440$. Similarly

$$\begin{aligned} P(Z > -1.00) &= P(-1.00 < Z < 0) + P(Z > 0) \\ &= P(0 < Z < 1.00) + P(Z > 0) \end{aligned}$$

(using symmetry again). From the table, this equals $.3413 + .5000 = .8413$.

It is my experience that trying to memorize the formulas that apply in each of the cases here is *not an efficient way to learn to carry out these calculations*. Instead, if you draw (or, with some practice, just visualize) a schematic picture of the standard normal p.d.f. and the locations of a and b, you should see how to apply symmetry and addition or subtraction to find the desired area $P(a < Z < b)$ in all of these cases.

We conclude with an example showing the full process of computing a normal probability using the Z-scores and the normal curve area table.

Example 10.28. Suppose that the weights of adult individuals of a certain species of fish are known to be normally distributed with mean $\mu = 500$ grams and standard deviation $\sigma = 30$ grams. We ask, if we catch a fish at random, what is the probability that its weight will be between 455 and 570 grams?

We begin by computing the Z-scores corresponding to the endpoints of the interval of weights we want:

$$y = 455 \longleftrightarrow z = \frac{455 - 500}{30} = -1.5$$

and

$$y = 570 \longleftrightarrow z = \frac{570 - 500}{30} \doteq 2.33.$$

From (10.20), we know that

$$P(455 < \text{weight} < 570) = P(-1.5 < Z < 2.33).$$

This is a case where the endpoints are on opposite sides of $z = 0$, so we split and use the symmetry:

$$\begin{aligned} P(-1.5 < Z < 2.33) &= P(-1.5 < Z < 0) + P(0 < Z < 2.33) \\ &= P(0 < Z < 1.5) + P(0 < Z < 2.33) \\ &\doteq .3531 + .4901 = .8432. \end{aligned}$$

In other words, there is about an 84% chance that an adult fish randomly chosen from the whole population will have a weight in that interval. $\triangle$

In case you are wondering what happens with the Z-score $Z = 0$, which apparently does not end up in either interval, rest assured. We could have included it in either interval or in both and the results would be exactly the same. The probability that a *continuous* random variable takes any single value is always *equal to zero*.

10.8. Chapter Project

The goals of this chapter's project (which will take several days) are:

- to introduce some basic features of the R statistics package (a much more extensive and powerful piece of software than the statistical routines in Excel – in fact R is probably the go-to "industry standard" software for academic statisticians at the moment),
- to see how R computes some descriptive statistics and graphics for understanding the "shape" or "distribution" of a data set, and
- to gain an intuitive understanding of the *Central Limit Theorem*, a basic result in statistics that underlies the wide applicability of normally-distributed random variables.

Background on R – Getting Started. The instructions here are geared toward the Windows version of R. Since R is a free, open source package, you can also obtain a copy to run on your own computer if you want, and there are Windows, Mac-OS, and Linux versions available. These have the same functionality but slightly different interfaces.

On PC's when you double-click the R shortcut from the standard desktop, you will launch the R-GUI window. Inside this is a workspace window with a white background. It *(only)* allows you to

- enter commands, and
- view numerical output.

It does not let you create integrated documents, and graphics are displayed in separate windows within the R-GUI, as you will see.

You can *print* the workspace window or any graphics window at any time. Just highlight that window and press the toolbar printer button.

To save your work, with comments and answers to questions (e.g. for the project report), I recommend

- opening another editor window (either Word or NotePad are fine for this on Windows systems), and
- using it to create a text file that will be your permanent record of your work.
- When you have a result you want to save from the R workspace, copy both the input command and the output and paste into the text file.
- Any time you want to add a comment to explain a calculation or answer a question, enter that directly into the text file.
- Save that text file in your personal network folder and print it when you want a hardcopy (from Word or NotePad).

A First R Session. In this first session, you will basically work through some examples illustrating features of R and some of the descriptive statistics we have talked about so far in the course. Keep a record of what you do for future reference.

In its most basic form, R provides a command-driven "statistical calculator" in which you type in data and commands to generate numerical and graphical output that are displayed. When you have the R window open, note the > input prompt after the opening message. This is where you should start.

(1) The following data set records systolic blood pressure readings from 15 adults (not a very healthy group, on the whole!) First we need to get the systolic blood pressure readings into R so that we can work with them. The basic way is to enter (after the $>$ prompt):

```
bpdata <-
c(172,140,123,130,115,148,108,129,137,161,123,152,133,128,142)
```

and press ENTER. Note: the $<-$ is formed by typing the $<$ and $-$ characters right next to each other. It is the *assignment operator* in R. This takes the data on the right and assigns it to the name on the left. The c, followed by the list of numbers, separated by commas is R's way of constructing *an ordered list or vector*. You should just get another input prompt if everything went OK with this (error messages if not). If you did get an error, try again, and try to follow the above exactly. You can verify that the numbers are entered correctly by typing in the name `bpdata` as a command; the numbers will be displayed in a row.

(a) We can compute the mean $\overline{y}$ and standard deviation s of the data in R (in several different ways!): Type in

```
ybar <- sum(bpdata)/length(bpdata)
mean(bpdata)
s <- sqrt(sum((bpdata - ybar)^2)/(length(bpdata) - 1))
sd(bpdata)
```

and press ENTER after each line to see the results. Note the way the sum function works in a very nice way here – in the computation of `ybar`, we only need to put in the name of the data list; the sum of all the elements is computed by default. Then in the computation of `s`, we can subtract *each entry* of the `bpdata` vector from `ybar` as shown! Of course we will usually want now to use the built-in functions `mean` and `sd`, but the first of each pair of computations shows how we can also use more primitive functions in R when we need them.

(b) Chebyshev's theorem mentioned in Chapter 9 is a result about the distribution of data sets. It says in a particular case that *there will always be at least* 75% *of the numbers in a data set within 2 standard deviations of the mean.* Let's use R to check that this is true here. In its most basic use, R can function as a simple numerical calculator(!)

```
upperb <- ybar + 2*s
lowerb <- ybar - 2*s
```

How many of the data values are in the range $[\overline{y} - 2s, \overline{y} + 2s]$, though? We could just count by hand, of course, with a data set this small. But R also contains a very rich collection of facilities for selecting, subsetting, and massaging data in numerous ways. Here is a basic way to select the portion of the data set in the range we want:

```
middle <- subset(bpdata, bpdata <= upperb & bpdata >= lowerb)
length(middle)/length(bpdata)
```

(This should be pretty self-explanatory; ask if you don't understand what this did!) Did Chebyshev's theorem "work" here?

(2) Now let's look at a different data set. This gives the number of parasites found in each of a sample of 100 individual foxes. The information is given there in table format:

Parasites	0	1	2	3	4	5	6	7	8
Foxes	69	17	6	3	1	2	1	0	1

So we want to create a data set containing 69 0 entries, 17 1 entries, etc. Clearly this is going to be tedious and error prone if we do it the most basic way(!) Fortunately, R lets us construct lists in more flexible ways besides just listing the entries. We can indicate how many times to repeat a given value if we want, like this:

```
foxpar <- rep(c(0,1,2,3,4,5,6,7,8),c(69,17,6,3,1,2,1,0,1))
```

The `rep` stands for "repetition" here. Note that the first list inside the `rep` gives the values of the number of parasites – the first row in the table format – and the second list gives number of times each value appears in the data – the second row. The first list could also be abbreviated as `0:8` (R's notation for a list of successive integer values).

(a) To construct a relative frequency histogram for the number of parasites per fox in R, let's begin by entering

```
hist(foxpar)
```

The histogram will be generated in a separate graphics window. This is a "default" version of the command. The `hist` command takes a number of options that control how the "bins" are selected and how the histogram is drawn as well. You will often want to modify the default results. For instance, looking at the output, how were the bins chosen here? Is that optimal for showing the different numbers of foxes with each number of parasites? Suppose we want bins centered on the integer values $0, 1, 2, 3, 4, 5, 6, 7, 8$ as we did in other examples. We can put in an option to make R do it that way too. The `breaks` option takes a list of values and makes those the boundaries of the bins for the histogram:

```
hist(foxpar,breaks=c(-0.5,0.5,1.5,2.5,3.5,4.5,5.5,6.5,7.5,8.5))
```

How is this different and why is it better than the default way that R drew the histogram?

(b) To practice using the commands we talked about before, calculate $\overline{y}$ and s for this data set. How many of the numbers were within two standard deviations of the mean in this case?

(c) The list of breaks for the histogram can be anything in R. What happens if you change the command above to

```
hist(foxpar,breaks=c(-0.5,0.5,1.5,2.5,3.5,4.5,5.5,8.5))
```

(deleting the 6.5 and 7.5 break points, so we're lumping together the counts for the values 6,7,8 parasites)? What is different about the new histogram that R drew in this case? Look at the graph carefully: do the heights of the boxes still represent *frequencies*? (Hint: what they do represent are *densities* that are computed by taking the frequency count and dividing by the width of the box in each case. What is true if we add the areas of the boxes drawn this way? We will see other reasons why this is a preferable way to draw histograms with unequal bin sizes later.) If we *really want frequencies*, we can make R do it that way:

```
hist(foxpar,breaks=c(-0.5,0.5,1.5,2.5,3.5,4.5,5.5,8.5),freq=TRUE)
```

However note that now R generates a warning message saying the AREAS in the plot are "wrong." Think about this. Why is the last box in this version of the histogram *misleading* if drawn this way?

Background on Probability Distributions in R. The R package contains built-in functions for computing probability functions for the discrete random variables we have discussed, plus functions for the probability density functions and cumulative distributions of the standard types of continuous random variables. Here's how it works. Each type of random variable is covered by a *family* of 4 functions distinguished by a *prefix letter* `d, p, q`, or `r`:

- `d` – the p.m.f. in the discrete case, or the p.d.f. in the continuous case,
- `p` – the c.d.f.
- `q` – the quantile function (we won't be using this one, but it can compute quartiles, deciles, percentiles, etc.)
- `r` – random number generator – for sampling random values of a random variable with that distribution

Following the prefix letter comes the rest of the function name and the inputs needed to compute the corresponding values.

Distributions of Sample Means. Work through the following example before and then follow the pattern here in the other parts of question A below. You do not need to hand in any work for this preliminary discussion, just for the parts of the question itself.

(1) Begin by generating a list of 100 random numbers from a particular *exponential* distribution (containing a parameter called the "rate" in R).[23] Use the command:

```
x <- rexp(100,1/3)
```

to assign the result to the name `x`. We will think of this as a *sample* of values from that distribution.

(2) Plot the density histogram for this data set `x` using the command

```
hist(x,freq=FALSE)
```

This should look roughly like the part of the graph of the p.d.f. for positive y, reflecting the fact that the random numbers are generated according to probabilities given by that density function.

(3) Our main goal is to understand what happens if we generate *lots* of such samples, find their sample means, then consider *the distributions of those sample means.*

(4) Here is one way to do this in `R` using a simple *for loop.* First we create space to store the means of 1000 different random samples:

```
means <- array(1:1000)
```

[23]This is a type of continuous random variable that we did not discuss in the text. You will see roughly what it's p.d.f. looks like below.

Next, we generate 1000 different random samples, find the means of each of them, and store them in the array created above:

```
for (i in 1:1000)
   {
   x <- rexp(100,1/3)
   means[i] <- mean(x)
   }
```

(Note: This could all be entered as one input line provided you leave spaces in the appropriate places. However, I think it is more readable (and it is easier to identify typos if you happen to make one) if you press ENTER at the end of each of the lines as you are typing this. If you do it that way, you should note that `R` generates a new input prompt `+` each time, indicating that you are still in the body of the `for` loop. The final } will close off the loop and take you back to the > prompt.)

(5) What does the distribution of the sample means look like? We can see that by generating a density histogram for the data in the `means` array:

```
hist(means,breaks=20,freq=FALSE)
```

This should look completely different from the histogram showing the distribution of the single sample above. In fact, what does this histogram remind you of?

(6) If you said, "a normal p.d.f.," good!! If you said something different, go back and look again.

(7) Now, which normal density is it? Well, by Theorem 10.26 from the text above, if we have 100 sampled values $Y_1, \ldots, Y_{100}$ from any distribution where $E(Y_i) = \mu$ for all i, and $V(Y_i) = \sigma^2$ for all i, and

$$\overline{Y} = \frac{1}{100}(Y_1 + \cdots + Y_{100})$$

is the sample mean, then we will have

$$E(\overline{Y}) = \mu$$

and

$$V(\overline{Y}) = \frac{\sigma^2}{100},$$

so the SD of the sample means will be $\frac{\sigma}{10}$. For the exponential random variables we are looking at, $\mu = 3$ and $\sigma = 9$. Hence the expected value of the sample means is also 3 and the SD of the sample means is .3.

(8) Let's overlay that normal density and see how well things match:

```
curve(dnorm(x,3,.3),add=TRUE)
```

(The `dnorm` is the normal p.d.f. the 3 is the μ and the .3 is the σ. You will use other normal densities in the questions below. They are specified in the same way.) This should match quite well (not perfectly, though!)

Assignment. Repeat the parts of the worked example below for the means of 1000 random samples of size 100 from each of the following distributions. Note: In each case to find the appropriate normal density curve, you will need to recall facts about expected values and variances of the Y_i individually, then use Theorem 10.26.

(A) A uniform distribution with $min = 4$ and $max = 10$. (These are the endpoints of the interval where the uniform density different from zero.) Use `runif(100,4,10)` to generate a random sample of size 100 from this uniform distribution. See Theorem 10.25 in the text for the appropriate μ and σ^2 in this case. Then apply Theorem 10.26 as in the example above.

(B) A Poisson distribution with $\lambda = 2$. From the discussion in the text, we know $\mu = \lambda$ and $\sigma^2 = \lambda$ for Poisson random variables. The appropriate R function to generate the samples of size 100 is `rpois(100,2)`.

What is happening in each case? Explain. Is the distribution of the sample means approximately normal every time?

Chapter Exercises

(1) Show that if all rolls of two distinguishable dice are equally likely, then the probability of having a 5 or a 6 on at least one of the dice is $5/9$. Hint: Count how many pairs (m, n) have $m = 5, 6$, or $n = 5, 6$ or both.

(2) In the text we used the fact that if $S = S_1 \cup S_2$ and $S_1 \cap S_2 = \emptyset$, then $P(A) = P(A \cap S_1) + P(A \cap S_2)$. Explain why the following more general statement should hold. Suppose $A = A_1 \cup A_2$, but we do not require $A_1 \cap A_2 = \emptyset$. Then

$$P(A) = P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

Hint: A Venn diagram of the union $A_1 \cup A_2$ showing a possibly nonempty $A_1 \cap A_2$ will explain what is going on here.

(3) In the situation of Example 10.8, what is the probability that all 5 surgeries are successful if $p = .9$? What if $p = .7$? What if $p = .5$?

(4) Compute the values $P(X = k)$ for all $0 \leq k \leq 10$ in Example 10.9 and compare with Figure 10.1.

(5) Suppose an urn contains $n_1 = 20$ red balls and 17 black balls. If $n = 5$ balls are chosen from the urn at random, what is the probability that exactly 2 of the balls are red,

(a) assuming the selection is done *with replacement*, then

(b) assuming the selection is done *without replacement.*

(6) We toss a fair die repeatedly.

(a) What is the probability the first 6 comes up on the 21st toss?

(b) What is the probability that first 6 comes up sometime on or after the 21st toss?

(7) An oil company will drill a succession of test wells in a given area until a productive well is found. Assume that the probability that any particular well will be productive is .3 and that the wells are independent of each other.

(a) What is the probability that the company will drill 9 wells before it finds the first productive one?

(b) What is the probability that exactly one well out of 10 wells drilled will be productive? Explain why your answers for part (a) and this part are different.

(8) Compute the values of the p.m.f. for a binomial random variable Y based on $n = 6$ trials with $p = .36$. Verify the statement that $E(Y) = np = 2.16$ by computing $E(Y)$ directly.

(9) (The binomial coefficients and "Pascal's triangle")

(a) Compute the numbers $\binom{n}{k}$ for $n = 0, 1, 2, 3, 4, 5$ and $0 \le k \le n$ in each case. Recall that $0! = 1$ by definition.

(b) Arrange your results with one row for each n and the values for $\binom{n}{k}$ in order from $k = 0$ on the left to $k = n$ on the right. You should notice something about how the numbers on the row for one n relate to the numbers for the next n (apart from the two 1's on the ends). Try to describe the pattern in words and in an equation.

(c) Show using (10.5) that for all n, k,

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

How does this relate to what you said in part (b)?

(10) The *probability mass function* (p.m.f.) of a discrete random variable with values in $\{0, 1, 2, 3, \ldots\}$ is defined as the function $f(k) = P(X = k)$.

(a) Explain why $f(k) \ge 0$ for all k.

(b) Explain why $\sum_{k=0}^{\infty} f(k)$ should equal 1. Note that when the values $f(k) = 0$ for all sufficiently large k (as is true for instance for binomial random variables), this is just a finite sum.

(c) Explain why the equation from part (b):

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = 1$$

is valid for a binomial p.m.f.

(11) Your front curb extends for 40 feet along the street. A delivery drone drops its cargo at a random point uniformly distributed along that 40 foot stretch.

(a) What is the probability that the drop point is between 5 feet and 15 feet along the curb from one endpoint?

(b) Give a "piece-wise" formula for the c.d.f. $F(x)$ for this random variable.

(c) What is the probability that the drop point is closer to the center than it is to either endpoint of the 40 foot stretch of curb?

(d) What is the probability that the drop point is within one SD of the midpoint? (See Theorem 10.25.)

(12) Consider the p.d.f. for a random variable Y given in Example 10.24.

(a) What is the probability $P(1.5 < Y < 2)$?

(b) What is the probability $P(Y > 2.5)$?

(13) Using the table in Figure 10.9, determine each of the following areas under the standard normal curve.

(a) $P(0 < Z < 1.43)$

(b) $P(0.72 < Z < 1.43)$

(c) $P(-2.17 < Z < -0.97)$

(d) $P(Z < 1.43)$

(e) $P(-0.62 < Z < 2.67)$

(14) The tail lengths of a species of lizard are normally distributed with mean $\mu = 75$ mm and SD $\sigma = 10$ cm.

(a) What is the probability that a randomly caught lizard of this species will have a tail length between 70 and 82 mm long?

(b) What is the probability that the tail length of a randomly caught lizard will be greater than 68 mm?

(c) If 10 of the lizards are caught independently, what is the probability that strictly fewer than 3 of them will have tail lengths greater than 68 mm? (Use your answer from part (b). How?)

(d) Suppose that 10 of the lizards are caught independently. What is the probability that the mean tail length for those 10 is greater than 68 mm? (Use Theorem 10.26 and your answer to part (b).)

(15) (Optional) This exercise requires some knowledge of limits. Let a be a real number at let r be a real number satisfying $-1 < r < 1$, or equivalently, $|r| < 1$. In this exercise, you will show that the infinite geometric series

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + ar^3 + \cdots \tag{10.21}$$

with first term a and ratio r between successive terms converges to (i.e. adds up to) $\frac{a}{1-r}$.

(a) Convergence of infinite series is defined using the finite *partial sums*, here

$$S_N = \sum_{n=0}^{N} ar^n = a + ar + ar^2 + \cdots + ar^N.$$

Show that this partial sum equals $\frac{a(1-r^{N+1})}{1-r}$. (Hint: Compute $S_N - rS_N$ and note the cancellations.)

(b) What is true about $\lim_{N\to\infty} r^{N+1}$ if $-1 < r < 1$?

(c) Use your answer to part (b) to show

$$\lim_{N\to\infty} S_N = \frac{a}{1-r}.$$

This is what is meant by saying that the infinite series (10.21) converges to $\frac{a}{1-r}$.

(d) Determine the sum of the infinite series

$$\sum_{k=0}^{\infty} \frac{3}{4^n}.$$

(e) Same question as part (d) for the infinite series

$$\sum_{k=0}^{\infty} \frac{(-1)^n 4^n}{5^n}.$$

(16) (Optional) This exercise requires knowledge of limits, and in particular the limit formula $\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x$.

(a) Start with the probability mass function for binomial random variable Y based on n trials and success probability p. Write $\lambda = np$ and consider

$$P(Y = k) = \binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Show that this can be rearranged to the form

$$\frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right).$$

(b) Now take the limit as $n \to \infty$ and use the limit formula given in the introduction to the problem to show that

$$\lim_{n\to\infty} P(Y = k) = \frac{\lambda^k}{k!}e^{-\lambda}.$$

This is the Poisson p.m.f. given in the text.

(17) (Optional) This exercise sketches a proof of the fact that the set of real numbers in the interval $[0, 1]$, namely the set of all real numbers x with $0 \le x \le 1$, is not a countably infinite set. The idea is known as the "Cantor diagonal argument." First recall that any real number with $0 \le x \le 1$ can be written as a decimal fraction of the form

$$x = .d_1 d_2 d_3 \cdots$$

where d_i are the decimal digits (all equal to $0, 1, 2, \ldots, 8$, or 9) of the number. Also recall that the decimal expansions of some rational numbers have either *terminating* or *infinite repeating* expansions. For instance

$$\frac{1}{2} = .5 = .4\overline{9} = .49999\cdots.$$

The proof will be by contradiction. Assume that *there is a one-to-one correspondence* between the natural numbers $\mathbb{N}$ and the numbers in $[0, 1]$ given, for instance, by

$$\begin{aligned} 1 &\longleftrightarrow .d_{11}d_{12}d_{13}\cdots \\ 2 &\longleftrightarrow .d_{21}d_{22}d_{23}\cdots \\ 3 &\longleftrightarrow .d_{31}d_{32}d_{33}\cdots \\ &\vdots \end{aligned} \tag{10.22}$$

In this listing, assume that the numbers with terminating decimal expansions are listed by their terminating forms with all zero digits after some point. Now consider the number with a decimal expansion $x = .e_1 e_2 e_3 \cdots$ that is obtained in this way: e_1 is any digit *different from* d_{11} *in the first number in* (10.22)*, and also different from* 9, then e_2 is any digit *different from* d_{22} *in the second number in* (10.22)*, and also different from* 9, then e_3 is any digit *different from* d_{33} *in the third number in* (10.22)*, and also different from* 9, and so forth. Explain why x cannot equal any of the numbers in (10.22). But on the other hand $0 \leq x \leq 1$ and (10.22) was supposed to be a complete list of all the numbers in $[0, 1]$. This contradiction shows no list like (10.22) can exist(!)

CHAPTER 11

Statistics of Sampling

11.1. Introduction

The idea of taking a random sample of likely voters and asking them their voting preferences in order to understand the possible outcome of an upcoming political election is possibly even over-familiar from recent Presidential and Congressional campaigns. The idea that the process of taking a collection of measurements in a scientific experiment is actually quite analogous to the process of selecting a sample from a population of potential voters might be less familiar. But as we will see in this chapter, this analogy is quite powerful and useful. The analogy is probably clearest when we are collecting measurements from a sample of organisms selected from population in the wild. We have seen a number of examples of datasets of this type in previous chapters. But even measurements in a lab experiment can be thought of as a sample from the "population" of all possible measurements of the quantity being measured.

In this chapter we will develop some of the theory underlying methods for making inferences about a population from the information contained in a sample. This theory will be applied in the next chapter to develop the framework of hypothesis testing using data.

11.2. Models of Sampling

Sampling to Estimate a Proportion. Suppose we have a population of N individuals and each one of them either has a certain characteristic or does not. Think, for example, of the characteristic that the individual will vote for the candidate of one specified party in an upcoming election. If n_1 individuals out of the total population have the characteristic and n_2 do not, so $n_1 + n_2 = N$, then by what we said in Chapter 10, the fractions

$$p = \frac{n_1}{N} \quad \text{and} \quad q = \frac{n_2}{N}$$

would represent the probabilities that a single individual chosen from the population at random will have the characteristic, or respectively, will not have the characteristic. The process of selecting a single individual and determining whether or not the characteristic is present is modeled by one of the random variables called *Bernoulli trials* in Chapter 10.[1]

The sampling done in public opinion polling, for instance, consists of choosing some sample of n individuals out of the total population of n and determining their

[1] We usually think the "success" outcome represents having the particular characteristic in this context.

voting preferences. Saying the sample is (perfectly) *random* means that every set of n individuals drawn from the population has (exactly) the same probability of being selected. As you might imagine this is not an easy condition to enforce or verify in practice. But it is a necessary condition for the validity of the methods we will discuss.

If m_1 out of the n individuals *in the sample* have the particular characteristic (e.g. intending to vote Democratic), then we would usually use the fraction $\frac{m_1}{n}$ as an *estimator* for $p = \frac{n_1}{N}$. Statisticians frequently use notation like this to indicate an *estimator* for the population parameter:

$$\boxed{\hat{p} = \tfrac{m_1}{n} \text{ is an estimator for } p = \tfrac{n_1}{N}.}$$

The most important thing to realize here is that, as we saw in the Chapter Project for Chapter 9, $\hat{p}$ *depends on which sample we take*. In other words, $\hat{p}$ is also *a random variable*, with an associated probability distribution for the possible values we would observe from repeated samples with that sample size n.

What is that probability distribution? The sampling process would almost always be equivalent to drawing n balls from an urn containing $N = n_1 + n_2$ balls *without replacement*: in the public opinion polling context, for instance, it makes no sense to allow the same person to be asked his or her opinion more than once. So, strictly speaking, we would would have a scaled version of a random variable with a *hypergeometric* probability distribution as given in (10.7) – "scaled" since the values are not just the numbers of successes, but rather the fractions $\frac{m_1}{n}$. But note that

$$P(m_1 \text{ successes}) = P\left(\hat{p} = \frac{m_1}{n}\right),$$

so we are in almost the same situation we described before.

However, as long as the sample size n is small relative to the population size N, then as we discussed in §2 of Chapter 10, a *binomial* probability distribution, corresponding to doing the selection process *with replacement*, would also be a good approximation. This approximation is typically used to understand probabilities in this situation because the binomial random variables are somewhat simpler to work with than the hypergeometric ones.

Since m_1 is the number of "successes" in the n Bernoulli trials, note that we can also describe the random variable $\hat{p}$ as the *mean*, or average, of the Bernoulli trials corresponding to the individuals in the sample. This means that we can use the result of Theorem 10.26 to determine the expected value and the variance or SD of $\hat{p}$. Since a single Bernoulli trial Y with success probability p has $E(Y) = p$ and $V(Y) = p - p^2 = pq$, and we are averaging n different Bernoulli trials with identical distributions, then as long as the trials are independent, we have

$$\boxed{E(\hat{p}) = p \quad \text{and} \quad V(\hat{p}) = \frac{pq}{n}.}$$

In this context, another name is usually used for the *SD of the estimator*. We call this the *standard error*. Hence here

$$\boxed{\text{standard error (of } \hat{p}) = SE = \sqrt{\frac{pq}{n}}.} \tag{11.1}$$

EXAMPLE 11.1. Results of political polling are usually represented by giving a value (usually as a percent), and an *interval of possible values around that*, for instance: "Our poll shows that 57%, plus or minus 3%, of likely voters favor candidate X." The estimate here would usually be derived from the $\hat{p}$ from the sample, and the interval would usually be the interval corresponding to ± 2 standard errors from the $\hat{p}$. This is an example of what is known as a *confidence interval* and the rationale for the choice ± 2 standard errors will become clear later.

A reasonable question to ask here is: How can one determine the standard error without knowing what p is? (And that's what the polling is supposed to do, after all, estimate the proportion p.) One possible approach is to use the value of $\hat{p}$ in the formula (11.1) to *estimate* the standard error. A somewhat more conservative method is to note that $pq = p(1-p)$ is maximized at $p = 1/2$, so no matter what p or $\hat{p}$ are, the standard error will always be *at most* $\frac{1/2}{\sqrt{n}}$. If $n = 1100$, for instance (a typical order of magnitude for a sample size in political polling), then we get

$$SE \le \frac{1/2}{\sqrt{1100}} \doteq .015.$$

Hence 2 SE's would be at most about .03, or three percent. In fact, from the range of possible values, it is always possible to estimate the size of the sample that was used to derive the interval if that is not reported.

Note also that our estimate for the standard error is strongly based on the assumption that the individual Bernoulli trials are *independent*. The results of political polling can be very unreliable if that is not true. $\triangle$

This sort of combination of a *point estimate* (a single numerical value) and an *interval estimate* (range of possible values) is a very common way to report estimated values of population parameters derived from data.

EXAMPLE 11.2. A random sample of $n = 300$ individuals are taken from a population of animals and $m_1 = 175$ of them are determined to be male. Let us give a 2-SE estimate for the interval of possible values of percentage of males in the whole population and indicate how this sort of interval estimate might be used.

We have $\hat{p} = \frac{175}{300} \doteq .583$, so our point estimate is that the males make up 58.3 percent of the population. Using the conservative method from the previous problem to estimate the standard error, we see

$$SE = \frac{1/2}{\sqrt{300}} \doteq 0.0289$$

So our 2-SE estimate would be $58.3\% \pm 5.8\%$, or 52.5% to 64.1%. Looking ahead, we will see later that since 50% is not in this interval, the value $p = .5$ for the actual proportion of males in the population is not very likely.

Another comment is in order here. Note that it would be conceivable to use the estimate $\hat{p} \doteq .583$ in the formula for the standard error. If we do that, then $1 - .583 = .417$ would be our estimate for q and

$$\sqrt{\frac{(.583)\cdot(.471)}{300}} \doteq 0.0285$$

is slightly smaller than our conservative estimate above. In practical terms, the slight difference indicates that both choices would yield very similar results. That is true in most cases. But the conservative method of taking $p = 1/2$ in the formula for the SE is usually considered to be preferable because it always applies and it will always yield an interval that is at least as large as the actual 2-SE interval. $\triangle$

Sampling to Estimate an Average. Now suppose we take a collection of measurements that yield numerical values in some interval of the real line (not just "yes/no" answers or one of two possible values that might be represented by numbers 1/0 as in the previous subsection). For instance, we might be thinking of taking very precise measurements of the heights of some sample of male first year college students to estimate the average height of a male first year college student. If $Y_1, \ldots, Y_n$ are the measured heights of the students in the sample, and μ represents the population average height, then one estimator for μ would clearly be the sample mean

$$\overline{Y} = \frac{1}{n}(Y_1 + \ldots + Y_n).$$

That is,

$$\boxed{\hat{\mu} = \overline{Y} \text{ is an estimator for } \mu.}$$

Once again, though, the $\overline{Y}$ is a random variable that has a certain probability distribution depending on which individuals are selected in the sample. We need to understand the probability distribution of $\overline{Y}$, and Theorem 10.26 contains the statements we need. That theorem says in this context that

$$\boxed{E(\overline{Y}) = \mu \quad \text{and} \quad V(\overline{Y}) = \frac{\sigma^2}{n},}$$

where σ is the population SD. As in the previous case, we call the SD of the estimator the *standard error*, or SE. In this case too, notice that exact value of the SE, namely

$$\boxed{SE \text{ (for } \overline{Y}) = \frac{\sigma}{\sqrt{n}}} \tag{11.2}$$

depends on information that we would not usually have at our disposal. Namely, σ, the population SD, is not something we would usually expect to know exactly. Almost always, the best we can do from data alone is to *estimate* the *population SD* using the *sample SD* from (9.6):

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}.$$

We reiterate that the sample SD is also an estimator – in this case, an estimator for the population SD. It also depends on the sample used and its values come with a particular probability distribution if the Y_i are normally distributed. We will see the underlying probability distribution for the closely related quantity

$$\frac{(n-1)S^2}{\sigma^2}$$

in a later section in this chapter.

Example 11.3. Refer to Table 1 in Chapter 9, which gives Height, Mouth, and Tube measurements for $n = 25$ pitcher plants from the *Darlingtonia californica* species. As we indicated in the Introduction, as long as we were reasonably sure that the measurements were independent values in a suitable sense (for instance, if we had taken the measurements at different locations for pitcher plants that were not interacting with each other), then it would be productive to view those measurements as a *random sample* from the whole population of this species of pitcher plants. We said in Example 9.10 that the Height measurements have sample mean $\overline{Y} \doteq 611.7$ and sample SD $S \doteq 265.1$. Hence our point estimator for the population mean height based on this sample is $\hat{\mu} = \overline{Y} \doteq 611.7$. The estimate for the SE would be

$$\frac{S}{\sqrt{25}} = \frac{265.1}{5} \doteq 53.02.$$

A 2-SE interval centered at the point estimator value would be

$$611.7 \pm 106.04$$

and this would represent an interval likely to contain the population mean height, μ. We will make this more precise in the next section. △

11.3. Sample Means from Normal Populations – t Distributions

The key point to understand about both types of estimators described in the previous section is that they are *random variables* with probability distributions derived from the probability distribution of the individual samples Y_i (the Bernoulli trials in the first case, and the individual measurements in the second). In this section, we will introduce the probability distribution that describes the way sample means vary when the Y_i are i.i.d. random variables that have *normal distributions*. This is a very important special case that turns out to be even more widely applicable than might be apparent at first. The idea here is:

THEOREM 11.4. *If Y_i are all normally distributed with expected value μ and variance σ^2, then*

(1) *The sample mean $\overline{Y}$ is also normally distributed with expected value μ and variance $\frac{\sigma^2}{n}$.*
(2) *More generally, any linear combination $a_1Y_1 + \cdots + a_nY_n$, where a_i are constant, is normally distributed with expected value $(a_1 + \cdots + a_n)\mu$ and variance $(a_1^2 + \cdots + a_n^2)\sigma$.*

This is a fact that is proved in more advanced mathematical statistics courses, but we will have to accept it without proof here. By Theorem 10.26, if each Y_i has expected value μ and SD σ, then

$$E(\overline{Y}) = \mu \quad \text{and} \quad V(\overline{Y}) = \frac{\sigma^2}{n},$$

as noted before. To compute the probability that $\overline{Y}$ is in a given interval, we would want to follow the standardization process from Theorem 10.27 and look at

$$Z = \frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\overline{Y} - \mu)}{\sigma},$$

which would have a standard normal distribution. However, as discussed in the previous section, we rarely have an exact value for σ and must usually "make do" with the sample SD, s instead. The resulting "approximate" standardized sample mean

$$T = \frac{\sqrt{n}(\overline{Y} - \mu)}{S} \tag{11.3}$$

does not have a standard normal distribution.

The exact probability distribution of the random variable T above was determined by William S. Gosset (1876 - 1937), a British statistician who worked for the Guiness brewery in Dublin. The brewery company did not allow its employees to publish results under their own names due to concerns about possible theft of trade secrets. In this case, Gosset was finally allowed to publish his statistical work under the pseudonym "Student" and the probability distributions he developed are still sometimes known as "Student's t-distributions," although the simpler name "t-distributions" is now more common. Gosset's main results here can be summarized as follows.

- The t-distributions depend on a parameter known as the *number of degrees of freedom*, ν,[2] which satisfies $\nu = n - 1$ in the case that T is given by (11.3) using the sample mean and SD of a sample of size n.
- For each n, there is an analytic formula for the p.d.f. describing the t-distribution, which has the form

$$f(t) = c\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2},$$

  where c is a constant, depending on ν, determined to make the total area between the graph of the p.d.f. and the t-axis equal to 1.
- The t-distribution p.d.f.'s have the same general shape as the standard normal p.d.f. In particular, they are all *symmetric* about $t = 0$.
- But their maxima are *lower* and they have *more area in the "tails"* as $t \to \pm\infty$ than the corresponding areas for the standard normal p.d.f. The graph in Figure 11.1 shows the standard normal p.d.f. in red and the p.d.f. for the t-distribution with $n - 1 = 9$ degrees of freedom in black.
- As ν increases, though, the p.d.f. for the t-distribution approaches the standard normal p.d.f. more and more closely.

We will say $Y \sim t(\nu)$ if Y has a t-distribution with ν degrees of freedom. (Note that the value of ν in (11.3) is $\nu = n - 1$ for means of samples of size n.)

THEOREM 11.5. *Let $Y_1, \ldots, Y_n$ be i.i.d. samples from a normal population with mean μ. Then*

$$T = \frac{\sqrt{n}(\overline{Y} - \mu)}{S}$$

has a $t(n-1)$ distribution.

The table in Figure 11.2 will be useful in working with the t-distributions.[3] Note that this table is laid out differently from the standard normal table in Figure 10.9 from Chapter 10. The number of degrees of freedom for the t-distribution is given in the column at the left. For each number ν of degrees of freedom up to 30 (and for another row giving the corresponding information for the standard normal), the row in the table gives a collection of values known as *percentage points* of the corresponding distribution. We will use the notation t_p for the p-percentage point, which

[2]This is the Greek letter "nu" – the equivalent of the Roman "n."

[3]From `http://www.dummies.com`, downloaded 11/22/2017. These and other probabilities for the t-distributions can also be computed using the `R` statistical package introduced in the Chapter Project for Chapter 10. With the free availability of software like `R`, the days of statistical tables are probably numbered(!)

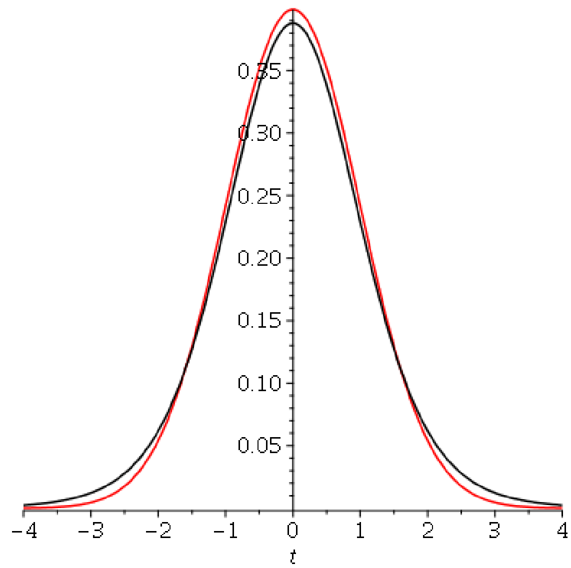


FIGURE 11.1. Standard normal and $t(9)$ densities compared.

by definition is the *lower endpoint* of the upper tail of the distribution for which $P(Y \geq t_p) = p$. The table gives t_p for the values $p = .4, .25, .1, .05, .025, .01, .005$, and $.0005$ and $Y \sim t(\nu)$ for the different ν shown.

EXAMPLE 11.6. For instance, check in Figure 11.2 to find the entry for $p = .05$ for $\nu = 12$ degrees of freedom. You should find 1.782288. This means that if $Y \sim t(12)$, then $P(Y \geq 1.782288) = .05$.

Similarly, for $\nu = 20$, $t_{.005} = 2.84534$. This means that $P(Y \geq 2.84534) = .005$ if $Y \sim t(20)$. $\triangle$

At this point we can begin to explain the rationale for the 2-SE intervals that we were using as examples in the previous section and indicate how the t-table in Figure 11.2 will allow us to do something analogous, but more precise and general, with sample means from normal populations.

From the standard normal table, Figure 10.9, we see that if Z has a standard normal distribution, then

$$P(-1.96 \leq Z \leq 1.96) = 2 \cdot 0.4750 = 0.9500.$$

This can also be stated in the following form. The number 1.96 is the .025-*percentage point for the standard normal distribution*, a fact which is often abbreviated by saying

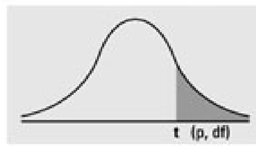
$$z_{.025} = 1.96.$$

This is true because $P(Z \geq 1.96) = .5 - .4750 = .025$.

If Y is any random variable with a normal distribution with mean μ and SD σ, then by Theorem 10.27, $Z = \frac{Y-\mu}{\sigma}$ has a standard normal distribution. It follows that

$$0.9500 = P\left(-1.96 \leq \frac{Y-\mu}{\sigma} \leq 1.96\right) = P(Y - 1.96\sigma \leq \mu \leq Y + 1.96\sigma).$$

Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).



df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%

FIGURE 11.2. A t-table.

The interval $Y \pm 1.96\sigma$ is called a 95% *confidence interval*. The precise meaning of this name is unfortunately rather easy to misunderstand. The confidence interval is based on a sample of values of Y and hence really should be thought of as a sort of "interval-valued random variable" – *we get different confidence intervals from different data* (i.e. different samples of values of Y). Saying this probability is equal to 0.9500 really means that *if we formed many intervals of this kind based on different sampled values of Y, then roughly 95% of those intervals would contain the μ – the expected value for the normal distribution describing the probability distribution of Y*.[4]

This is the basis for a rule of thumb that if Y has a normal distribution (or even an approximately normal distribution), then the probability that Y is within about 2 SD's of its mean is about 0.95, or 95%. In other words, the 2-SE intervals we were looking at before were based on saying 1.96 is very close to the "round number" 2 and using the standard errors for the estimators we were considering.

But now we can also see why the t-table in Figure 11.2 is set up the way it is. Namely, the information there is presented in exactly the form needed to derive a confidence interval for the population mean of a normally-distributed population based on the information contained in a sample of size n from the population, in the form of the sample mean, $\overline{Y}$, and the sample SD, S. We see that if $p/2$ is any one of the probability values given in the top row of the t-table, if $T \sim t(\nu)$, then by the symmetry of all of the $t(\nu)$-p.d.f.'s,

$$1 - p = P(-t_{p/2}(\nu) < T < t_{p/2}(\nu)).$$

Hence, if $T = \frac{\sqrt{n}(\overline{Y}-\mu)}{S}$ as in Theorem 11.5, then

$$1 - p = P\left(-t_{p/2}(n-1) \le \frac{\sqrt{n}(\overline{Y}-\mu)}{S} \le t_{p/2}(n-1)\right),$$

which (after some algebraic rearrangement) becomes

$$1 - p = P\left(\overline{Y} - t_{p/2}(n-1)\cdot\frac{S}{\sqrt{n}} \le \mu \le \overline{Y} + t_{p/2}(n-1)\cdot\frac{S}{\sqrt{n}}\right).$$

This gives the following general formula for the endpoints of a $(1-p)\times 100\%$-confidence interval for the population mean of a normally-distributed population:

$$\boxed{\overline{Y} \pm t_{p/2}(n-1)\cdot\frac{S}{\sqrt{n}},} \tag{11.4}$$

based on the information contained in sampled values $Y_1, \ldots, Y_n$, under the assumption that the samples are independent and identically distributed (i.i.d.).

Example 11.7. Let us return to the pitcher plant height data from Example 11.3. We will assume here that the heights follow a normal distribution. If we wanted a 90%-confidence interval, for instance, then we would take $p = .1$ and hence $p/2 = .05$. Since there were $n = 25$ sampled values used to compute the values of $\overline{Y} \doteq 611.7$ and $S \doteq 265.1$ given before, we use the t-table with $\nu = 25 - 1 = 24$

---

[4] It is easy to fall into the trap of saying "there is a 95% chance that the interval we get from the formula contains μ." However, that does not really make sense because for every such interval, μ is either contained in the interval or it is not. This is a subtle point about what we are doing here that will require careful attention in order to avoid misconceptions.

degrees of freedom. The value $t_{.05}(24) = 1.710822$, so the endpoints of the 90% confidence interval are

$$611.7 \pm 1.710822 \cdot \frac{265.1}{5} \doteq 611.7 \pm 90.7$$

On the other hand, for a 98% confidence interval we would take $p = .02$, so $p/2 = .01$ and then from the table $t_{.01}(24) = 2.49126$ and the endpoints would be

$$611.7 \pm 2.49126 \cdot \frac{265.1}{5} \doteq 611.7 \pm 132.1.$$

As we should expect, the 98% confidence interval from a given collection of samples will be wider than the 90% confidence interval because we are saying we want a recipe for making intervals more of which will contain the population mean. Since the $\frac{S}{\sqrt{n}}$ factor in (11.4) also depends on how spread out the sample values are, though, we cannot say that a 90% confidence interval will *necessarily always* be narrower than a 98% confidence interval based on a sample of a fixed size from the same population, even though that *usually* would be true. △

11.4. The Central Limit Theorem and an Application

In the Chapter Project for Chapter 10, we observed that even if sampled values $Y_1, \ldots, Y_n$ come from a probability distribution that is far from being normal, the *sample means* $\overline{Y} = \frac{1}{n}(Y_1 + \cdots + Y_n)$ have a probability distribution that seems to be close to normal if the sample size n is sufficiently large. The way normal distributions appear from the averaging process is the subject of a fundamental result in mathematical statistics known as the *Central Limit Theorem*, or CLT for short. Here is one of the simpler versions of the theorem. This has been vastly generalized by statisticians, but the simpler version will suffice for our purposes.

Theorem 11.8 (Central Limit Theorem). *Let the $Y_1, \ldots, Y_n$ be i.i.d. random variables with finite expected value μ and SD σ. Then the probability distribution of*

$$\frac{\sqrt{n}(\overline{Y} - \mu)}{\sigma}$$

will be close to a standard normal distribution if n is sufficiently large.

Some comments are in order:

- The hypotheses that μ and σ are finite are not automatically true, but they will hold in all of the cases we look at.[5]
- A commonly-used rule of thumb is that "sufficiently large" typically means that the agreement is already quite close for $n \geq 30$ or so.

[5]Technical note: This condition is a reasonably mild one; all of the examples we have seen satisfy it. But there are some continuous random variables with p.d.f.'s $f(y)$ for which the improper integrals

$$E(Y^k) = \int_{-\infty}^{\infty} y^k f(y)\, dy$$

with $k = 1$ and/or $k = 2$, which go into computing μ and σ^2, do not converge .

- The agreement also gets *closer the larger n is*, and a more precise statement is that the distribution of
$$U_n = \frac{\sqrt{n}(\overline{Y} - \mu)}{\sigma}$$
actually *converges to the standard normal distribution* as $n \to \infty$ in a certain sense.
- Intuitively, the statement of the theorem also means that $\overline{Y}$ itself will have an approximately normal distribution with mean μ and SD equal to $\frac{\sigma}{\sqrt{n}}$. The form given in the theorem corresponds to applying the standardization formula from Theorem 10.27.
- We will not try to give a formal justification of this fact because it requires additional techniques that are beyond the scope of this text.[6] Instead, we will simply rely on the intuition we gained from the Chapter Project in Chapter 10.

The CLT basically says that (under some relatively mild conditions) any random variable obtained by an *averaging process* using enough values will have an approximately normal distribution, *no matter what the underlying probability distribution of the individual sampled values is.* This says, for instance, that the methods for deriving confidence intervals from the last section would also apply if our data consisted of the *sample means* from samples with sufficiently large sample size n. The CLT, or perhaps more general statements of this kind that do not require such strict hypotheses on identical distribution of the Y_i, is also often offered as the underlying reason why so many quantities observed in the natural world (e.g. heights and weights of plant or animal bodies, etc. which are effectively results of a sort of averaging of many genetic and environmental factors) follow normal distributions.

Our main example in this section will concern the binomial random variables from Chapter 10. Recall that a binomial experiment consists of some number n of i.i.d. Bernoulli trials Y_i. The corresponding binomial random variable gives the number of successes in those trials, so we have
$$Y = Y_1 + \cdots + Y_n = n \cdot \frac{1}{n}(Y_1 + \cdots + Y_n).$$
Since the Y_i have finite expected value p and variance $pq = p - p^2$, the CLT applies. We see that if n is sufficiently large, the scaled binomial $\frac{1}{n}Y$ is the mean of the Bernoulli trials, and the CLT implies that it will have an approximately normal distribution with expected value p and SD equal to $\sqrt{\frac{pq}{n}}$. Multiplying by n to get Y itself, we have the following Corollary of the CLT.

Proposition 11.9. *If n is sufficiently large, then a binomial random variable Y based on n trials with success p has approximately the same distribution as a normal random variable U with expected value np and SD equal to $\sqrt{npq}$.*

The probability histogram in Figure 11.3, which is copied from Figure 10.1 for your convenience, shows an example. Even for $n = 10$ (not as large as the rule of thumb $n = 30$ discussed above), and $p = .4516$, the shape of the probability

[6]Technical note: The book *Mathematical Statistics with Applications* by Wackerly, Mendenhall, and Scheaffer gives a somewhat more precise version of the theorem and discusses one proof by the technique of moment generating functions.

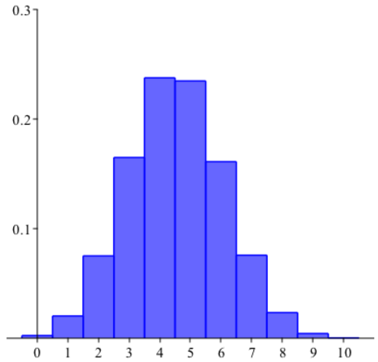


FIGURE 11.3. Binomial probabilities, $n = 10$, $p \doteq .4516$.

histogram for the binomial is already quite similar to the shape of the region under the graph of a normal p.d.f. .

Proposition 11.9 is the basis of a traditional method for *approximating binomial probabilities using the standard normal table*. The only (small) "wrinkle" here is that the binomial is a *discrete* random variable while the normal is *continuous*.

This means that (as is suggested by the histogram in Figure 11.3), in doing this approximation, we want to think of the value $Y = k$ for the binomial as corresponding not to a single value of the normal,[7] but rather *to the whole interval of values from* $U = k - 1/2$ *to* $U = k + 1/2$. If we want the probability that Y has one of several possible values, we just take a union of such intervals. Here is an example.

EXAMPLE 11.10. Let Y be a binomial random variable based on $n = 30$ trials with $p = 0.6$. Let us use the idea sketched above to approximate the binomial probability

$$P(Y = 17, 18, \text{ or } 19).$$

By Proposition 11.9 we want to use a normal random variable U with expected value $np = 30 \cdot 0.6 = 18$ and SD $\sqrt{30 \cdot 0.6 \cdot 0.4} \doteq 2.68328$. The binomial probability will be approximated by $P(16.5 \le U \le 19.5)$, where we have approximated $P(Y = 17)$ by $P(16.5 \le U \le 17.5)$, and so forth. Standardizing and using the standard normal table as in the examples from Chapter 10, we have

$$\begin{aligned} P(16.5 \le U \le 19.5) &= P\left(\frac{16.5 - 18}{2.68328} \le Z \le \frac{19.5 - 18}{2.68328}\right) \\ &\doteq P(-0.56 \le Z \le .56) \\ &= 2P(0 \le Z \le .56) \\ &= 2 \cdot .2123 \\ &= .4246. \end{aligned}$$

[7]As for any continuous random variable and one particular value, $P(U = 17) = 0$.

The exact value is

$$P(Y = 17, 18, \text{ or } 19) = \binom{30}{17}(.4)^{17}(.6)^{13} + \binom{30}{18}(.4)^{18}(.6)^{12} + \binom{30}{19}(.4)^{19}(.6)^{11}$$
$$\doteq .42303,$$

so the approximation is good to two decimal places. That would be sufficient for many purposes.

The awkwardness of computing the binomial coefficients such as

$$\binom{30}{17} = 119,759,850$$

and raising p and q to such large powers *by hand* probably sufficiently explains the appeal of this normal approximation to binomial probabilities. However, it must be said that such calculations using the exact binomial p.m.f. formula are not at all difficult with the software and other computational resources at our disposal at the present time. △

11.5. χ^2 and F Distributions

In this section we will rapidly[8] introduce two other types of continuous random variables that are used in various ways connected with estimating population *variances* and developing the hypothesis tests that we will discuss in the next chapter. The first of these are the so-called χ^2 random variables.[9] As for the t-distributed random variables introduced earlier in this chapter, there is a whole family of different χ^2 random variables parametrized by a positive integer ν called the *number of degrees of freedom*. We will write $\chi^2(\nu)$ to represent the χ^2 distribution with ν degrees of freedom.

We say Y has a $\chi^2(1)$ distribution if Y is the *square of a standard normal*, that is $Y = Z^2$, where Z is a normally-distributed random variable with expected value $\mu = 0$ and variance $\sigma^2 = 1$. Probabilities for these can be computed using the standard normal table.

Example 11.11. Let $Y \sim \chi^2(1)$. Then for any $a > 0$,

$$P(0 \le Y \le a) = P(0 \le Z^2 \le a) = P(-\sqrt{a} \le Z \le \sqrt{a})$$

where Z is a standard normal. For instance,

$$P(0 \le Y \le 2) = P(-\sqrt{2} \le Z \le \sqrt{2}) \doteq P(-1.41 \le Z \le 1.41) = 2 \cdot .4207 = .8414,$$

using the standard normal table from Figure 10.9. △

Then the other $\chi^2(\nu)$-distributed random variables for $\nu > 1$ are described as follows.

Definition 11.12. Y has a $\chi^2(\nu)$ distribution if $Y = Z_1^2 + \cdots + Z_\nu^2$, where the Z_i are independent standard normals.

[8] A euphemism for "sketchily"(!)

[9] χ is the Greek letter "chi" and χ^2 is usually read as "chi square."

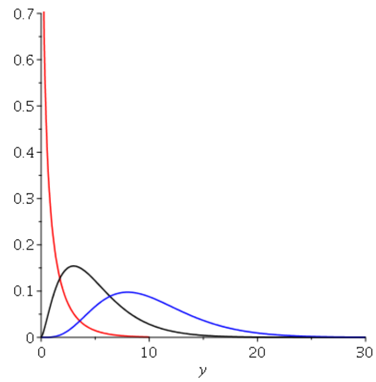


FIGURE 11.4. $\chi^2(\nu)$ p.d.f.'s – $\nu = 1$ in red, $\nu = 5$ in black and $\nu = 10$ in blue.

There is an analytic formula for the p.d.f. of a $\chi^2(\nu)$ random variable which has the form

$$f(y) = \begin{cases} cy^{(\nu/2)-1}e^{-y/2} & y > 0 \\ 0 & y \leq 0. \end{cases} \tag{11.5}$$

Note $Z_i^2 < 0$ is never true so we expect the p.d.f. to be equal to 0 for all $y < 0$. The constant c, depending on ν, is chosen to make the total area under the graph of the p.d.f. for $y > 0$ equal to 1. Figure 11.4 shows the p.d.f.'s for $\chi^2(1)$ in red, for $\chi^2(5)$ in black, and for $\chi^2(10)$ in blue.

The connection between χ^2-distributed random variables and estimators for variance comes from the form of the sample variance:

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}.$$

From this, if the Y_i are samples from a normal population with SD equal to σ, we get

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{Y_i - \overline{Y}}{\sigma}\right)^2. \tag{11.6}$$

Now, if the $\overline{Y}$ was replaced by μ in each term on the right side of this equation, then we would be summing the squares of n independent standard normals. But since we have the sample mean there rather than the population mean, the $Y_i - \overline{Y}$ are *not independent*. Note that their sum is always zero, so any $n-1$ of the terms determines the last one. This is the genesis of the idea of "$n-1$ *degrees of freedom*," in fact. Only $n-1$ one of those terms are independent, so we might say the sum on the right side of (11.6) has only $n-1$ "degrees of freedom." Our next statement is not entirely obvious from what we have said, but it can be proved with some fairly sophisticated algebraic manipulation.

THEOREM 11.13. *Let Y_i be i.i.d. samples with a normal distribution with variance σ^2. Then*

$$V = \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{Y_i - \overline{Y}}{\sigma} \right)^2$$

has a $\chi^2(n-1)$ distribution. Moreover, $\overline{Y}$ and S are independent random variables.[10]

Tables of percentage points for the $\chi^2(\nu)$ distributions analogous to the table of percentage points for the $t(\nu)$ distributions given in Figure 11.2 have been compiled and are available in many places. We will give one such table in the next chapter when we discuss hypothesis testing using χ^2-distributions. Software such as `R` also gives ways to find any such percentage points that are required.

The final topic for this chapter is to introduce yet another family of distributions derived from the $\chi^2(\nu)$ family.

DEFINITION 11.14. *Let $V_1 \sim \chi^2(\nu_1)$ and $V_2 \sim \chi^2(\nu_2)$ be two independent χ^2 random variables. The quotient*

$$F = \frac{V_1/\nu_1}{V_2/\nu_2}$$

is said to have an $F(\nu_1, \nu_2)$-distribution.

The number ν_1 is called the number of degrees of freedom in the numerator and the number ν_2 is called the number of degrees of freedom in the denominator. Tables of percentage points for the $F(\nu_1, \nu_2)$ distributions analogous to the table of percentage points for the $t(\nu)$ distributions given in Figure 11.2 and to those for the $\chi^2(\nu)$ distribution have also been compiled and are available in many places. Software such as `R` also allows one to compute them easily.

As you might expect if you think about the formula (11.6), and the formula for S obtained from it:

$$S^2 = \sigma^2 \cdot \frac{V}{n-1}$$

[10]Technical Note: There is also a very close connection between these formulas and the one used before to define the $t(n-1)$ distribution. Consider the approximate standardization

$$T = \frac{\sqrt{n}(\overline{Y} - \mu)}{S}$$

using S. If we solve for S in the equation in Theorem 11.13, then

$$S = \sigma \cdot \sqrt{\frac{V}{n-1}}.$$

Hence the formula for T above can be rewritten by substituting for S:

$$T = \frac{\left(\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \right)}{\sqrt{\frac{V}{n-1}}}.$$

The numerator $\left(\dfrac{\overline{Y} - \mu}{\sigma/\sqrt{n}} \right)$ is the standardization of the normal random variable $\overline{Y}$; the denominator is the square root of a $\chi^2(n-1)$ random variable V, divided by its number of degrees of freedom. The numerator and denominator are independent by the last statement in Theorem 11.13. The quotient is then a random variable with a $t(n-1)$-distribution. Indeed, this is sometimes taken as the *definition* of a $t(n-1)$-distributed random variable.

(in which V has a $\chi^2(n-1)$ distribution), the values of the F above will be used to compare the sample variances S_1^2 and S_2^2 of two groups of samples and determine whether it is reasonable to assume that they come from populations with *same* population variance σ^2. This is a part of a test for whether there is evidence to say two population means are different or not based on the information in groups of samples from each of the two populations.

11.6. Chapter Project

The goals of this chapter project are:

- To gain some more intuitive understanding of meaning of confidence intervals by developing a graphical demonstration. The concept of confidence intervals is a notoriously "slippery" one. It is easy to go from the intuitively appealing "confidence level" $(1-p) \times 100\%$ as in the text to statements that seem to be saying the same thing, but are misleading at best, and completely meaningless at worst. To use this idea reliably to make inferences from real-world data, it is very important to understand *exactly* what it means.
- To gain some experience with user-developed `R` functions.

A Confidence Interval Demo. If we have $Y_1, \ldots, Y_n$ in the "large sample" case ($n \geq 30$ or so) so the CLT has "kicked in," then it follows from what we said in the text that an the endpoints of an (approximate) $(1-p) \times 100\%$ confidence interval for the population mean μ is produced by the formula:

$$\overline{Y} \pm z_{p/2} \frac{S}{\sqrt{n}},$$

where $z_{p/2}$ is the percentage point for the standard normal (the same as a number in the last row of the t-table in Figure 11.2). But what does this really mean, and what is the right way to interpret the "confidence level" $(1-p) \times 100\%$?

The following `R` code will generate a collection of *one hundred* 95% confidence intervals based on samples of size $n = 50$ from a normal population with mean $\mu = 10$ and $\sigma = 2$. It produces two things:

(a) a graphic window display – each interval is drawn with a horizontal line, and the location of the population mean is marked by the vertical line
(b) the number of intervals containing μ, printed in the R console window as the output of the function.

```
confdemo <- function()
{
  endpts <- matrix(1:200,nrow=100,byrow=T)
  nin <- 0
  for (i in 1:100)
  {
     ysamp <- rnorm(50,10,2)
     ybar <- mean(ysamp)
     S <- sd(ysamp)
     endpts[i,1] <- ybar - qnorm(.975,0,1)*S/sqrt(50)
```

```
    endpts[i,2] <- ybar + qnorm(.975,0,1)*S/sqrt(50)
    if ((endpts[i,1] <= 10) && (10 <= endpts[i,2]))
      nin <- nin + 1
  }
  plot(c(endpts[1:100,1],endpts[1:100,2]),c(1:100,1:100),
    xlab="confidence intervals",ylab="number of trial")
  for (i in 1:100) lines(c(endpts[i,1],endpts[i,2]),c(i,i))
  lines(c(10,10),c(1,100))
  return(nin)
}
```

Enter the code line by line. When you get no error messages, you call the function with an input command:

```
confdemo()
```

This will generate the graphical display and the output to the R console.

Questions.

1. Look over the `R` code carefully and try to figure out what each line does. Explain. In particular, how are the plotting commands (`plot` and `lines`) working? Note: The `c` is the `R` constructor for a list. `c(a,b)` denotes a list with entries a, b.
2. Use the `confdemo` function to generate the plot showing one hundred 95% confidence intervals for the mean, generated from samples of size $n = 50$ from a normal population with $\mu = 10$ and $\sigma = 2$.
3. Explain what your plot shows. In particular, how many of the intervals contain $\mu = 10$ and how many do not? How does that relate to the 95% confidence level?
4. Will the plot *always* show exactly 95 intervals containing the population mean $\mu = 10$? Try calling the function several more times if you got exactly 95 the first time. Does it make sense to say 95% of the confidence intervals produced here contain $\mu = 10$? In what sense?
5. An `R` function can also take *input values* that are used when the function is called. The names go in the set of parentheses in the first line, and then you supply values for them when you call the function. Modify the `confdemo` function so that you can specify the `sample size` n, and the confidence level $1-p$. **Hints:** The first line of your new function definition can look like

   ```
   confdemo2 <- function(sampsize,conflevel)
   ```

   and every where you used the sample size $n = 50$ before, now you will put in the name of the input `sampsize` instead. Similarly everywhere that used the $p/2 = .025$ from before, now you will want `conflevel/2` (this one is a bit trickier – you will need to recall what the `qnorm` function does). If you like, you can really jazz this up so that the μ and σ can be specified as well.
6. Suppose you now use `confdemo` to study the 95% confidence intervals for the mean, generated from samples of the new size $n = 70$ from the same

normal population with $\mu = 10$ and $\sigma = 2$. What will change (apart from the fact that the computation will take somewhat longer)? When you check your intuition, be sure to look carefully at the horizontal axis scales.

7. Suppose you now use `confdemo` to study the 99% confidence intervals for the mean, generated from samples of size $n = 50$ from the same normal population with $\mu = 10$ and $\sigma = 2$. What will change? When you check your intuition, be sure to look carefully at the horizontal axis scales.
8. A common *misconception* about confidence intervals can be stated as follows: "When you compute the 95% confidence interval for the mean from a particular sample, there's a 95% chance that the population mean is contained in your interval." This statement (taken literally) is actually *meaningless* – why? What could you do to modify the statement so that it makes sense and is a true statement about confidence intervals?
9. Another common *misconception* about confidence intervals can be stated as follows: "When you increase the sample size n, the width of the $(1 - p) \times 100\%$ confidence interval always decreases." This statement is actually false – why? What could you do to modify the statement so that it makes sense and is a true statement about confidence intervals?
10. A final common (and tempting) *misconception* about confidence intervals deals with the following situation. Say we are using the interval to decide whether evidence from samples supports the hypothesis that a population mean μ has a particular value μ_0. If the value μ_0 is contained in a confidence interval but is far from the midpoint (even very close to one endpoint), it is tempting to think that the evidence indicates possibly $\mu \neq \mu_0$. This is an *incorrect* deduction in this situation. Look carefully back at your results. Explain why concluding $\mu \neq \mu_0$ on the basis of one confidence interval where μ_0 is close to an endpoint is *not a correct conclusion.*

Assignment. Lab report containing input, output, and answers to the questions.

Chapter Exercises

(1) Discuss how you would find or estimate the following standard errors.
  (a) The standard error for estimating the proportion if we have a sample size $n = 400$, and $m_1 = 321$ out of the 400 have the specified characteristic. Discuss how you are determining the SE.
  (b) The standard error for estimating a population mean with a sample size $n = 20$ from a normal population with $\sigma = 2.1$. What if the population mean is not known, but your sample has SD $S = 1.9$?

(2) Following what we said in Example 11.1, how big a sample size would we need to take to get a plus or minus 2-SE interval of $\pm 1\%$ around the $\hat{p}$ estimate for a proportion? (Use the conservative estimate for the standard error.)

(3) Find the 2-SE interval estimates for the Mouth and Tube data from Table 1 in Chapter 9.

(4) Find the 95% and 98% confidence intervals for the Mouth and then the Tube data from Table 1 in Chapter 9. Follow the method illustrated in Example 11.7 in the text.

(5) Assume the data of trout Length and Weight measurements given in Table 3 of Chapter 5 come from a random sample of the population of trout in some area.

   (a) Determine 90% and 95% confidence intervals for the population average length.

   (b) Determine 90% and 95% confidence intervals for the population average weight.

(6) Use the method of normal approximations to binomial probabilities discussed in Example 11.10 to estimate the following binomial probabilities. Compare with the exact values computed using `R` or other suitable sofware.

   (a) $P(Y = 12)$ for a binomial random variable Y with $n = 35$ and $p = .27$.

   (b) $P(10 \leq Y \leq 15)$ for a binomial random variable Y with $n = 27$ and $p = .46$.

   (c) $P(Y > 20)$ for a binomial random variable Y with $n = 50$ and $p = .64$.

(7) Suppose a random sample of 2000 voters are asked their voting preference in an upcoming election. Of those sampled, 1213 favor candidate A. Give 95% and 98% confidence intervals for the proportion of voters favoring A in the whole population of voters. Use the percentage points from the standard normal (the last row in the t-table).

(8) This exercise discusses conditions under which we would expect the normal approximation to a binomial probability to be reasonably close. This often happens for n much smaller than the $n \geq 30$ we would expect from the CLT. The idea is that we would expect the normal approximation to relatively good if a 3-SE interval for p is entirely contained in the interval $(0, 1)$, that is

$$0 < p - 3\sqrt{\frac{pq}{n}} \quad \text{and} \quad p + 3\sqrt{\frac{pq}{n}} < 1.$$

(If this were not true, then a significant portion of the area under a normal curve approximating the binomial probability histogram would fall outside the range of possible values for p.)

   (a) By manipulating the inequality algebraically, show that $0 < p - 3\sqrt{\frac{pq}{n}}$ is true exactly when $n > 9\left(\frac{p}{q}\right)$.

   (b) Similarly, show that $p + 3\sqrt{\frac{pq}{n}} < 1$ is true exactly when $n > \left(\frac{q}{p}\right)$.

The two parts of this exercise can be combined to give the rule of thumb that we expect the normal approximation to the binomial to yield good results whenever

$$n > 9\left(\frac{\text{larger of } p, q}{\text{smaller of } p, q}\right)$$

(9) The median age of residents of a certain country is 29 years. What is the probability that at least 55 out of a random sample of $n = 100$ residents will be under the age of 29 years?

(10) (For readers who know calculus)

(a) How would you determine the value of c in (11.5) to make the total area under the graph of $f(y)$ equal to 1 in case ν is even?

(b) Determine the values of c when $\nu = 2, 4, 6$.

(c) When ν is odd, we need to introduce an important special function called the gamma-function to describe the values of c:

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, dt.$$

Show that this function satisfies the functional equation $\Gamma(x+1) = x\Gamma(x)$ and $\Gamma(1) = 1$, so $\Gamma(x) = (x-1)!$ when x is an integer.

(d) Show

$$c = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}$$

for all integers $\nu \geq 1$. The value $\Gamma(1/2) = \sqrt{\pi}$. Knowing that and the functional equation from part(c), explain how to determine the value of c for all positive integers.

CHAPTER 12

Hypothesis Testing and Statistical Inference

12.1. Introduction

For the purposes of this introductory discussion, imagine first that you are a scientist. You have collected some data from an experiment and you believe that it shows that a certain pattern exists in the real world situation you are studying. For instance, you might think you have shown that a particular non-till farming method reduces soil erosion as compared with conventional farming methods. *How do you demonstrate that the measurements you have made provide evidence for claiming that that pattern actually exists? Your arguments should be in a form that can be checked and evaluated by other scientists, and your goal is essentially to convince those others that your claims are true, using some accepted mathematical tools.*

Similarly, now suppose you are a public opinion pollster, and you think you have found some interesting pattern in the answers to a question on a survey you have designed. For instance, this pattern could be that a particular political candidate seems to be leading in the run-up to an election.[1] Again, *how do you demonstrate that this pattern is there, in a way that is verifiable and that will convince others of the correctness of your claims?*

In both cases, you would need to be able to show convincingly that whatever you saw in the data you collected was not simply the result of some naturally occurring variation or some random glitch in the measurement process. In both cases, the issue is whether the same pattern observed in a *sample* (that is, the measurements made) is also true of the whole population from which the samples were drawn (that is, all possible instances of using the farming method, or all the people in the population from which the respondents of the questionnaire were chosen).

Key Idea: The typical methods used by most scientists would include performing a statistical analysis on the data aimed at showing that *the observed results would be extremely rare if the claimed pattern did not actually hold true.*

If that can be done successfully, the fact that the pattern in the measurements made was observed can be due to one of two possible reasons: Either it is

- due to the fact that the pattern is *truly there* in the whole population, or else,
- due to some truly "bad luck" in how your experiment turned out.

In the second case, you "hit the jackpot," but in a negative way. You observed a very rare event that does not agree with the true state of things.

[1]As we said in Chapter 11, this might be a situation you are sick of hearing about. But it is an important application with significant practical implications.

This probably sounds rather convoluted at first, but it is the logic behind the *hypothesis tests* we will discuss, and it is the most important idea behind this part of our course.

12.2. The Logic of Hypothesis Tests

A fundamental assumption underlies all of the methods we will describe in this chapter, namely that whatever measurements we are analyzing come from a *random sample* of the population of all possible measurements of the type we are considering. If that condition is violated, these methods can yield misleading results. However, techniques used in the design of experiments to ensure reasonable compliance with the requirement of randomness are somewhat beyond the scope of our discussion and we will not concentrate on that side of the equation.

Null and Alternative Hypotheses. Assuming we have collected data that yields something like a truly random sample from the population of possible measurements, the way the testing process is described typically involves two *competing explanations* for the results of the experiment:

- A *"null hypothesis"* (often denoted H_0 if we need a symbolic abbreviation) that says essentially "there's nothing there – the results were just the result of naturally occurring variation or some random glitch in the measurement process." Of course, the exact statement is usually more precise than this – it usually involves a more specific statement about the thing(s) you are measuring.
- An *alternative hypothesis* (denoted H_a) that is is some assertion that a claimed pattern is really there. This is also usually stated in more precise form(!)

The statistical null and alternative hypotheses can be, and usually are, different from the scientific hypotheses that were made in designing the experiment that produced the data. These statistical hypotheses are proposed explanations for the data that was observed. The goal is to decide whether the weight of the evidence contained in the measurements supports H_a, or whether that evidence could reasonably be explained by H_0. From the point of view of a researcher, of course, it is typically the alternative hypothesis H_a that is the "preferred" alternative. Not being able to rule out the null hypothesis is usually taken as a negative outcome, since in that case we are saying the results we saw could just be due to chance.

Here are two examples that should make the distinction between the null and alternative hypotheses clearer.

Example 12.1. A random sample of $n = 50$ compact fluorescent light bulbs was chosen from the production line of the manufacturer and the mean weight of mercury per bulb was measured to be 5.3 mg, with an SD of .5 mg. Assume the mercury amounts in all such bulbs are normally distributed. Looking at the 5.3, it might be tempting to say: "The average amount of mercury in these bulbs is > 5 mg." But does the data support that? Might the particular sample chosen have just contained especially mercury-heavy bulbs, not typical of the whole population? In this case the null hypothesis could be H_0 : the actual population average mercury level μ satisfies $\mu \leq 5$ mg per bulb. And the alternative hypothesis could be: H_a : the actual population average mercury level is $\mu > 5$ mg per bulb. $\triangle$

EXAMPLE 12.2. *StarLink* (a registered trademark) corn is a genetically engineered variety that was approved as a source of feed for animals, but never approved for human consumption. A study by the USDA shows that 99 out of 1100 samples of corn taken from US (human) corn-based food products were contaminated by traces of StarLink corn. At the same time the corresponding agency of the Mexican government does a similar study and finds that 100 out of 1200 samples of corn-based food products taken from Mexican manufacturers contain traces of StarLink. Note that the contaminated proportions are $\widehat{p_{US}} = 99/1100 = .09$, while $\widehat{p_M} = 100/1200 \doteq .083$ is smaller. Does the data support the conclusion that StarLink contamination of human food products is different in the US than in Mexico? Here we could have $H_0 : p_{US} = p_M$ (the actual proportions in the two countries are just the same and the difference observed in the samples is due to chance). The corresponding alternative hypothesis might be $H_a : p_{US} \neq p_M$. It would also be possible to use $H_a : p_{US} > p_M$ if the goal is to decide if the evidence shows that the proportion in the US is actually greater. $\triangle$

The conclusions from statistical hypothesis tests are usually phrased in very cautious language. One usually says something like "there is enough evidence to reject the null hypothesis" rather than saying "the alternative hypothesis is definitely true." Similarly a negative result might be described by saying "there is not enough evidence to reject the null hypothesis" rather than saying "the alternative hypothesis is definitely false." The point here, of course, is that the results from one experiment, no matter how strong or weak, can only provide limited evidence one way or the other. In addition, the scientific hypotheses underlying the experiment or the statistical forms H_0 and H_a are always subject to revision if further evidence indicates that previous thinking was incorrect.

Type I and Type II Errors, α and β. A statistical test aimed at choosing between the two hypotheses H_0, H_a can "go wrong" in two different ways:

- *Type I Error:* We could reject H_0 (and implicitly assert the evidence favors accepting H_a) when H_0 is actually true.
- *Type II Error:* We could fail to reject H_0 when it is actually false.

Both types of errors are of concern, but the Type I error is usually considered to be, if anything, more serious. That is because making an incorrect conclusion when an apparent pattern is due only to chance variation can throw off subsequent research, can lead to inappropriate recommendations for real-world action, and can have other undesirable consequences. Making a Type II error, on the other hand, means essentially that we *missed* a pattern that is there (perhaps by being too cautious in assessing the data). A common view would be that since there is always the chance of *catching* the pattern with another experiment later, Type II errors are, in a sense, easier to correct. If there were some urgency or time pressure involved in the real-world situation under study, a Type II error could also lead to serious real-world consequences, though. For instance, consider a study of the effectiveness of a vaccine for an infectious disease (informally: H_0 : no benefit from the vaccine and H_a : there is a benefit). If the testing was being done while an epidemic was in progress, then a Type II error leading to a decision not to use the vaccine might lead to loss of lives that could have been saved by using the vaccine. In real life, in fact, people are often willing to try unproven medical treatments in situations where they have nothing to lose, or in which any chance of a positive outcome, even a small one, outweighs other risks.

In any case, it is important to realize that most statistical tests are set up *to make the chance of a Type I error small.* This is exactly the *Key Idea* stated at the beginning of this chapter, restated in a more precise form. The chance of making a Type I error is usually denoted by the Greek letter α ("alpha"), and typical values for α used in the design of statistical hypothesis tests are .05, or .01, or perhaps even smaller values. Saying $\alpha = .05$, for instance, says that we want to set up our test so that if H_0 is actually true, then roughly 95 times out of 100 the results of the test will correctly indicate that H_0 *should not* be rejected. Or equivalently, Type I errors would happen roughly only 5 out of 100 times when H_0 is true. In other words, as we said before, *results indicating that we should reject H_0 would be extremely rare if the claimed pattern did not actually hold true (that is if H_0 is actually true)* .

The chance of making a Type II error is denoted by another Greek letter, β ("beta"). In a sense, the real quantity of interest here is $1 - \beta$, the chance of making a correct conclusion and rejecting H_0 when H_0 is false. This is called the *power* of the test, and ideally we would like the power to be as close to 1 as possible. But it is important to realize that the power, and equivalently, the value of β, is usually somewhat *harder to control* than α because they will typically depend on characteristics of the population from which the measurements are being taken that may not be known exactly. For instance, if H_0 is the hypothesis that a population mean is equal to some particular number μ_0, then H_0 being false means the population mean is something different from μ_0. The value of β will usually depend on the exact value of that population mean. Thus, the power of the test is actually *a function of μ = true population mean*, not just a constant. Statisticians have developed techniques for studying β and the power of tests as well, but they are somewhat beyond the scope of our treatment of this subject. Without going into the details, we can just say that in order to obtain a test with a given desired power, we would typically have to be able to choose a sample size n that was sufficiently large.

Since making measurements in carrying out experiments will usually incur real-world costs in one fashion or another, this might not always be feasible. For example, a polling organization might not have the time, the manpower, or the access to contact information to carry out a phone survey of a random sample of $n = 5000$ likely voters within a few days before an election in order in a test of voter preferences, even though having n that large might be necessary to to achieve a small β value.

Test Statistics, Rejection Regions. Here is the overall plan for a typical statistical test:

- Using assumptions about the distribution of the possible measurement values from the population, some *test statistic* that has a known probability distribution under the assumption that the null hypothesis H_0 is *true* is identified.
- Some desired Type I error probability α is selected. (Often $\alpha = .05$ is used as a standard choice; smaller values might be used too. Values $\alpha > .05$ would almost never be used in practice since there is something of a consensus that Type I error probabilities that large are unacceptable.)

- Using that probability distribution, a *rejection region* is identified. This is a range R of possible values of the test statistic with the property that, under the assumption H_0 is true, the chance that the test statistic lies in R is α.
- Then starting from the observed measurements, we compute the value of the test statistic, determine whether it lies in the rejection region or not.
- If, so we say "there is evidence indicating H_0 should be rejected" or something similar; if not we say "there is not enough evidence to reject H_0."

This probably seems somewhat abstract without a specific example to think about. Hence we will proceed immediately to first example tests, the "z-tests for means and proportions."

12.3. Tests for Means and Proportions

A t-test for a mean ("large sample case"). Suppose we have made a relatively large number $n \geq 30$ of numerical measurements of a particular characteristic from randomly chosen individuals in some fixed population (think something like lengths of individual adult fish from some particular species). Call those measurements $Y_1, \ldots, Y_n$. Then it is reasonable to expect that the sample mean

$$\overline{Y} = \frac{Y_1 + \cdots + Y_n}{n}$$

should give a good estimate of the population average μ. From the Central Limit Theorem, the intuition gained in our Chapter Projects, and the discussion from Chapter 11, recall that we expect that the values of $\overline{Y}$ from different samples should be normally distributed, and that the conversion to the standard normal should go by this z-score formula:

$$\frac{\overline{Y} - \mu}{\sigma/\sqrt{n}},$$

where σ is the *population* SD. Of course we don't know the exact value of σ in practice, so we would usually estimate that using $S =$ the sample SD. Moreover, we don't know the value of μ either.

Let's consider the following situation. Suppose we have a particular "candidate" value $\mu = \mu_0$ in mind as part of a null hypothesis– $H_0 : \mu = \mu_0$. Moreover, suppose the data seems to indicate, say, that $\mu > \mu_0$. That will be the alternative hypothesis H_a.

Under the assumption that H_0 is true, we have seen in Chapter 11 that the test statistic

$$T = \frac{\overline{Y} - \mu_0}{S/\sqrt{n}}$$

has a $t(n-1)$-distribution. For common values of α the percentage points of the t-distributions from Figure 11.2 show how to set up a rejection region. See Example 12.3 below for an example. For $n > 30$, though, the $t(n-1)$ distribution is very close to the standard normal. So we can use the standard normal table rather than the t-table to find the associated percentage point. For instance, from the standard normal table, the chance that $T > 1.65$ would be about .05. This would gives the rejection region for a test with $\alpha = .05$ – we reject H_0 whenever $T > 1.645$ and we do not reject it otherwise.

EXAMPLE 12.3. Now let's illustrate how this would be applied. Suppose we had measured data that looked like this ($n = 30$ measurements):

$$5.2, 5.3, 5.9, 8.8, 8.9, 7.1, 5.9, 6.3, 7.2, 5.0,$$
$$5.3, 4.9, 3.2, 5.8, 6.2, 6.0, 7.2, 7.2, 4.2, 5.1,$$
$$6.8, 7.8, 6.2, 5.2, 5.5, 4.3, 4.7, 4.7, 6.7, 6.3$$

We have $\overline{Y} = 5.96$ and $S = 1.30$ (rounding to two decimal places). For our example test, let's take:

- $H_0 : \mu = \mu_0 = 5.75$ and
- $H_a : \mu > \mu_0 = 5.75$.

Then, we compute:

$$T = \frac{\overline{Y} - \mu_0}{S/\sqrt{n}} = \frac{5.96 - 5.75}{1.30/\sqrt{30}} = .8677$$

From the t-table, we find $t_{.05}(29) \doteq 1.6973$. We still use the t-table here because with $n = 30$ our test statistic has a $t(29)$-distribution. Since $.8677 < 1.697$, we *cannot reject* H_0 on the basis of the evidence in this data.

The numbers above are a hypothetical dataset generated from a normal distribution with true mean $\mu = 6.4$ and true $\sigma = 1.2$(!) As you should have noticed in the Chapter Project on sampling, there is a lot of potential variability in sampling or measuring. In this case, the sample mean $\overline{Y}$ is especially small relative to the actual μ, and the sample SD S is larger than the actual σ. Both of these contributed to a test statistic T that was "smaller than expected." This is a case where we are making a Type II error! Although we won't discuss how this would be derived, the power of this test with $\mu = 6.4$ would be about $1 - \beta = .83$–that is, the probability of a Type II error would be about $\beta = 1 - .83 = .17$. The above dataset "beat the odds" in a way, but that sort of thing would happen roughly 17% of the time. △

There are a number of related forms of tests for means based on the form of the alternative hypothesis. For a test with $\alpha = .05$, for instance, and $n > 30$,

- if $H_a : \mu > \mu_0$ (an "upper-tail" test), then we would reject H_0 if $T > 1.645$
- if $H_a : \mu < \mu_0$ (a "lower-tail" test), then we would reject H_0 if $Z < -1.645$
- if $H_a : \mu \neq \mu_0$ (a "two-tail" test), then we would reject H_0 if $Z < -1.96$ or $Z > 1.96$.

These rejection regions all come from the areas of regions under the standard normal curve. Note that the area between 0 and 1.96 is approximately .4750, so the area in the upper tail from 1.96 to $+\infty$ is about .025. Similarly the area in the lower tail from $-\infty$ to -1.96 is also about .025 by the symmetry of the standard normal curve. This means that the rejection region for the two-tail test has total area about $.025 + .025 = .05$ and the chance that T lands in that region under the assumption that $H_0 : \mu = \mu_0$ is true is $\alpha = .05$ (approximately).

Hypothesis Tests and Confidence Intervals. As you may suspect from this discussion, there is very close connection between the rejection region for a two-tail hypothesis test and a confidence interval for the mean as discussed in Chapter 11. Recall that we have seen that in this large sample case, the endpoints of the 95% confidence interval for μ would be computed from the sample mean $\overline{Y}$ and $S =$ the sample SD by the formula

$$\mu = \overline{Y} \pm 1.96 \cdot \frac{S}{\sqrt{n}}$$

If we add the assumption that $\mu = \mu_0$ from the null hypothesis, when you look at this formula, it is not too difficult to see that it is saying the following. The values of $\overline{Y}$ for which we would *not* reject $H_0 : \mu = \mu_0$ are exactly the $\overline{Y}$ such that

$$\mu_0 - 1.96 \cdot \frac{S}{\sqrt{n}} < \overline{Y} < \mu_0 + 1.96 \cdot \frac{S}{\sqrt{n}}$$

(since it is exactly those values of $\overline{Y}$ for which $T = \frac{\overline{Y}-\mu_0}{S/\sqrt{n}}$ satisfies $-1.96 < Z < 1.96$).

In other words, the confidence interval gives the range of "believable" values for μ based on the the mean and SD of the sample. The rejection region for the test is something like the *complement of (that is, the part of the number line outside)* the interval computed using the confidence interval formulas, but using μ_0 from the null hypothesis rather than the sample mean from the data.

If we only have $n \leq 30$ measurements, then the actual properties of the t-distribution for $\nu = n - 1$ degrees of freedom must be taken into account. We will see what to do in those "small sample" cases later.

***Z*-tests for a Proportion.** Say we have asked a random sample of n people from a particular population a "yes-or-no" question and some number Y of them answer "yes." Suppose P is the proportion of the whole population who would answer "yes" if asked and we want to estimate this proportion. From the sample, as discussed in Chapter 11, we can estimate P using $\widehat{P} = Y/n$.

The theoretical basis for the test in this case is the following mathematical result: Under $H_0 : P = P_0$,

a) when n is relatively large (the cutoff value $n > 30$ is often used), *or more generally*
b) when nP_0 and $n(1-P_0)$ are both relatively large (a typical "rule of thumb" is both are ≥ 5),

then it can be shown[2] that the test statistic

$$Z = \frac{\widehat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$$

has an approximately standard normal distribution.

So we can set up hypothesis tests with $\alpha = .05$ using the same rejection regions as above in the case of a large-sample t-test for a mean:

- If $H_a : P > P_0$ (an "upper-tail" test), then we would reject H_0 if $Z > 1.645$
- If $H_a : P < P_0$ (a "lower-tail" test), then we would reject H_0 if $Z < -1.645$
- If $H_a : P \neq P_0$ (a "two-tail" test), then we would reject H_0 if $Z < -1.96$ or $Z > 1.96$.

[2]Technical note: This uses the same reasoning based on the Central Limit Theorem that we used in Chapter 11 to derive the normal approximation to a binomial distribution.

EXAMPLE 12.4. Supppose that $n = 1000$ likely voters are surveyed and $Y = 550$ of them say they will vote for candidate Jones in the next city council election. With $\alpha = .05$, is there sufficient evidence to say that Jones will win the coming election? We take $H_0 : P = P_0 = .5$ and $H_a : P > .5$. (Note that this is one case where the statistical null hypothesis is somewhat different than we might expect – Jones will only lose the election if $P < .5$, and the boundary case $P = .5$ would be a "dead heat" in the election.) In any case, with the values of n, Y as given, we have $n > 30$ and $nP_0 = 500 = n(1 - P_0)$, so both conditions a) and b) above are satisfied and we can use this approach. We compute

$$Z = \frac{.55 - .50}{\sqrt{\frac{(.5)(.5)}{1000}}} \doteq 3.16 > 1.645.$$

This is strong evidence to indicate that we should reject H_0. Note that any value $P_0 < .5$ would yield an even larger Z. So taking $P_0 = .5$ in this case is justified since it is the boundary case between a loss for Jones and a win for Jones. $\triangle$

Large-Sample Tests for Differences of Means and Differences of Proportions. There are similar tests of hypotheses about the difference of two population means or two proportions. The ones we will discuss are valid *only in the large sample case* (both groups of samples of size at least 30). The analysis is based on the assumption that the samples are random and independent.

Say we have measured some quantity in a random sample of size $n_1 \geq 30$ from a first group, and measured the same quantity in a random sample of size $n_2 \geq 30$ from a second group. (The two numbers n_1 and n_2 can be different, but both should be in the large sample range.) Call the two groups of measurements $Y_1, \ldots, Y_{n_1}$ and $X_1, \ldots, X_{n_2}$. So we have two sample means $\overline{Y}$ and $\overline{X}$, as well as sample SD's S_1 (from the Y_i) and S_2 (from the X_j).

Say μ_1 and μ_2 are the population means of the two groups. To test $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$, we would use the test statistic:

$$Z = \frac{\overline{Y} - \overline{X}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

since under the assumption $\mu_1 = \mu_2$, Z has (approximately) a standard normal distribution. The rejection region would be set up according to the value of α exactly as before.

Similarly, for a test on a difference of proportions, say we have asked the same "yes-or-no" question to random samples from two different groups. Say Y_1 out of $n_1 \geq 30$ in the first group say "yes" while Y_2 out of $n_2 \geq 30$ in the second group say "yes." (There is also a more general "rule of thumb" parallel to b) in the discussion of the one-proportion test above that is sometimes used; we will not discuss that, however.) We take $\widehat{P_1} = Y_1/n_1$ estimating P_1 and $\widehat{P_2} = Y_2/n_2$ estimating P_2 (the two population proportions). Then to test $H_0 : P_1 = P_2$ versus $H_a : P_1 \neq P_2$, we need some estimate of the common value under the null hypothesis in order to construct our test statistic. The one that is commonly used is a "pooled" estimator of the common proportion (assuming the null hypothesis):

$$P_p = \frac{Y_1 + Y_2}{n_1 + n_2}.$$

Then the test statistic is

$$Z = \frac{\widehat{P_1} - \widehat{P_2}}{\sqrt{\widehat{P_p}(1 - \widehat{P_p})} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

and the rejection regions are as before.

Note that the "yes-or-no" question can be anything with only those two possible answers. This applies much more generally than to only the polling situation. For instance,

EXAMPLE 12.5. Refer back to Example 12.2 above. Is there evidence to show that different proportions of corn products contain traces of StarLink in the US and in Mexico? Note that the contaminated proportions are $\widehat{p_{US}} = 99/1100 = .09$, while $\widehat{p_M} = 100/1200 \doteq .083$. So $P_p = \frac{199}{2300} = .087$. We have

$$Z = \frac{.09 - .83}{\sqrt{(.087)(.913)}\sqrt{\frac{1}{1100} + \frac{1}{1200}}} \doteq .595$$

For the two-tail test with $\alpha = .05$, we would be looking for $Z > 1.96$. So this is not strong enough evidence to reject H_0. The proportions in the US and in Mexico could be the same and the difference could be due to random variation. $\triangle$

T-tests for a Mean ("small sample case"). When the number of measurements n is smaller than 30, then the tests for the mean presented above use the same test statistic, but must use different rejection regions coming from the appropriate "Student's t-distributions" discussed in Chapter 11. Another important thing to note in this situation is that we cannot assume that the Central Limit Theorem has "kicked in" either, so *it is necessary to assume, in addition, that the measurements are drawn from a population described by a normal distribution* in this case.

When $n < 30$, we use the same test-statistic as before:

$$T = \frac{\overline{Y} - \mu_0}{S/\sqrt{n}}.$$

But now the rejection regions are found from the t-table and will depend on n:

- If $H_a : \mu > \mu_0$ (an "upper-tail" test), then we would reject H_0 if $T > t_{.05}$ for $n - 1$ degrees of freedom
- If $H_a : \mu < \mu_0$ (a "lower-tail" test), then we would reject H_0 if $T < -t_{.05}$ for $n - 1$ degrees of freedom
- If $H_a : \mu \neq \mu_0$ (a "two-tail" test), then we would reject H_0 if $T < -t_{.025}$ or $T > t_{.025}$ for $n - 1$ degrees of freedom.

The needed information to set up the two-tail rejection region for $\alpha = .05$ is given in Figure 11.2 in Chapter 11. For example, if we had $n = 10$ sample measurements, then the rejection region for an upper-tail test with $\alpha = .05$ would be $T > 1.833$ from the $\nu = 9$ row of the table. (The t-distribution involved is the one that has 9 degrees of freedom – note that the number of degrees of freedom is always $n - 1$.) The corresponding two-tail test would have rejection region $T < -2.262$ or $T > 2.262$. Note that these are more restrictive than the corresponding rejection regions for the large-sample test. This is because the distribution of $\overline{Y}$ and the test statistic T *has even more variability when n is small.* So the rejection region must be smaller than the corresponding rejection region for a large-sample test with the same $\alpha = .05$.

EXAMPLE 12.6. A random sample of ten 100 km^2 plots in a forest are chosen and the number of robin nests in each area is counted, yielding the following data:

$$310, 311, 412, 368, 447, 376, 303, 410, 365, 350$$

The sample mean is $\overline{Y} = 365.2$ and the sample SD is $S = 48.417$. Is there sufficient evidence to claim that the average number of robin nests per 100 km^2 in this forest is less than 380? We compute

$$T = \frac{365.2 - 380}{48.417/\sqrt{10}} \doteq -.967$$

Since $T > -1.833 = -t_{.05}$ for 9 degrees of freedom, we *do not* have sufficient evidence to reject $H_0 : \mu = 380$ with $\alpha = .05$. $\triangle$

12.4. p-VALUES AND HYPOTHESIS TESTS

When reporting the results of a statistical hypothesis test, it is common to provide an additional (or sometimes an alternative) piece of information called the *p-value* of the test. Another name is the *attained significance level.* The reason for that name is the fact that the Type I error probability α is often called the *significance level* of the test. So, for instance if we have decided on $\alpha = .05$ and the result of the test is to reject H_0, we might say that the result is "significant at the .05-level." (This is the precise meaning of statements like "the results of the experiment were statistically significant.")

Look back at Example 12.4. There we had a test where the test statistic had the value $Z = 3.17$. We could say "there is evidence to reject H_0 at the $\alpha = .05$ significance level, since $3.17 > z_{.05}$. But in fact since $Z = 3.17$ is quite a bit larger than $z_{.05} = 1.645$, in a way we are *understating the strength of our evidence* if we stop there.

DEFINITION 12.7. The attained significance level p is *the smallest value of α for which H_0 would be rejected with the observed value of the test statistic.*

In Example 12.4, since the area under the standard normal curve to the right of 3.17 is approximately .00076, using an upper-tail test, we would reject H_0 with this data for any $\alpha \geq .00076$. We say the p-value is $p = .00076$.

An alternative way to think about what the p-value of a test is telling us is this: the p-value is the *chance of observing the given value of the test statistic, or something "more extreme," if H_0 is true.* So *very small* values of p indicate *very strong evidence for rejecting* H_0 according to the Key Idea at the start of this Chapter. Conversely, larger values of p (often any $p > .05$) is taken as indicating the evidence is too weak to reject H_0. Since there is so much arbitrariness in the $\alpha = .05$ value, nowadays, in fact, it is actually common just to report the p-value of a test and leave the interpretation of whether to accept or reject H_0 up to the reader(!)

Example 12.8. Refer back to Example 12.1. Since $n = 50$, we use the large-sample formulas based on the standard normal curve. Our test statistic is

$$T = \frac{5.3 - 5}{.5/\sqrt{50}} \doteq 4.24$$

For the upper-tail z-test of $H_0 : \mu = 5$ versus $H_a : \mu > 5$, we have a p-value $p = 1.1 \times 10^{-5} = .000011$. This would be interpreted as very strong evidence to reject H_0. Note that it is the sample size $n = 50$ that is making it come out this way. If we knew the population SD was .5, under the null hypothesis a single measurement of 5.3 would have a z-score of

$$\frac{5.3 - 5}{.5} = .6$$

This is not very large (less than one SD above 5), so nothing remarkable. But the fact that we have the average of $n = 50$ measurements being 5.3 means that $T = .6 \times \sqrt{50} \doteq 4.24$. $\triangle$

Statistical Significance and Practical Significance. Some caution is certainly a good thing in applying the methodology of hypothesis testing. As we indicated before, the significance level of a test α (or the p-value) is based entirely on the probability of rejecting the null hypothesis when it is actually true. So we are really only taking the Type I errors into account. For that reason, the statistical significance level of a test is only an imperfect measure of how much information we get from the test.

Moreover, there are certainly times when a statistically very significant test result (say one that gives a very small p-value) has little *practical* significance. For example, a regression analysis on a large data set[3] might yield the result that the slope of the regression line was nonzero with an attained significance level of $p = .0001$. But the corresponding confidence interval of values for the slope could contain only very small numbers (e.g. .01 to .03). There might not be much of a practical difference between

- saying a unit change in x produces a change of between .01 and .03 in y, or
- saying a unit change in x produces no change in y

if the values of y are in the 1000's, for instance. Statistical significance and real-world practical significance are not the same things!

12.5. Small-sample Test for Differences of Means

There are tests for differences of population means in the small sample case as well. However, the assumptions required to apply the tests in their basic forms are *much* more restrictive than in the large sample case:

- We must assume that the two sets of samples come from normal populations (since we may not be able apply the Central Limit Theorem to claim normality for the distribution of the sample means otherwise), and
- We must assume that the *variances of the two normal distributions are equal*.

[3]See Section 6 below for a more detailed description of what this means.

We will only discuss how to determine whether the basic form is appropriate and how to carry out the test in that case. For the more sophisticated methods needed to deal with non-normal data and/or the unequal variance case, we suggest that the reader refer to more advanced statistics texts, or perhaps even better, consult an expert statistician.

First we note that there are a number of statistical tests for determining whether it is reasonable to assume that given data is sampled from a normal distribution or not. One of the most intuitive, visual, and qualitative methods is known as the *normal probability plot*, or *QQ plot*. The normal probability plot is a scatter plot showing a comparison between a given data set and a reference normal distribution.[4] The closer the normal probability plot comes to being linear, the closer the agreement between the data and what we would expect from samples from a normal distribution. This could be quantified using the R^2 statistic or in other ways. But in fact, "eyeballing" the normal probability plot is often sufficient for our purposes.

Other, more quantitative tests are also available, including the Shapiro-Wilk W-test. The Shapiro-Wilk test is a formal test of

- a null hypothesis H_0: The data are sampled from a population that is normally distributed, versus
- an alternative hypothesis H_a: The data are sampled from a population that is not normally distributed.

The exact form of the test statistic and how rejection regions and p-values are determined are beyond the scope of this discussion, but the conceptual framework is exactly what we are discussing in this chapter.

However, we wish to caution the reader that we think most experienced statisticians would say that *tests such as the Shapiro-Wilk test are probably of little use* for basic data analysis where we need to assess whether tests based on normality assumptions are appropriate. The reason is that for small sample sizes, the test will often not distinguish between data drawn from normal distributions and data from other distributions. On the other hand, for large sample sizes,[5] even very small deviations from normality (for instance using a large sample from a t-distribution with a large number of degrees of freedom) will result in very small p-values indicating that the normality assumption should be rejected. Since a $t(100)$ distribution is for all practical purposes equivalent to a standard normal, this is an inappropriately sensitive tool to apply to this sort of question.

Software for generating normal probability plots (and also for carrying out the Shapiro-Wilk test) is included in the `R` statistical package and we will see examples of how both methods are implemented there in the Chapter Project.

Assuming that there is no reason to suspect the two datasets do not come from normal distributions, the next question to consider is whether the equal variance assumption is reasonable. This is where the F-distributions considered in section

---

[4]Technical note: The name "QQ plot" stands for *quantile-quantile plot* and this indicates how the plot is formed. The normal probability plot effectively sorts the data and then compares that sorted list with the expected values of the order statistics–the minimum, the next smallest, the third smallest, etc., up to the maximum–of a sample of the same size drawn from a standard normal distribution.

[5]Technical note: The `shapiro.test` function in `R` limits the size of the input data list to $n = 5000$ to avoid this issue.

5 of Chapter 11 come into their own. Looking back at Theorem 11.13 and Definition 11.14, we can see that if S_1^2 and S_2^2 are the sample variances of two sets of samples of sizes n_1 and n_2, drawn from normal populations *with equal variances*, then the statistic

$$\boxed{F = \frac{S_1^2}{S_2^2}} \tag{12.1}$$

will have an F-distribution with $n_1 - 1$ degrees of freedom in the numerator and $n_2 - 1$ degrees of freedom in the denominator. Hence, we have the makings of a test of

- H_0: the population variances are equal, versus
- H_a: the population variances are different.

The test statistic is F above. With a sufficiently detailed table of percentage points for the F-distributions or software such as `R` we can construct a 2-tail rejection region and carry out a test with any given Type I error probability α.

Example 12.9. (*Note:* Hypothetical data is used in this example.) A state Wildlife Conservation Department carried out a study to understand the build-up of chemical residues in the brain tissue of birds of a particular species. Random samples of $n_1 = 11$ adult birds and $n_2 = 9$ juveniles produced mean brain tissue DDT levels of $\overline{Y_1} = .039$ for the adults and $\overline{Y_2} = .024$ for the juveniles (in units of parts per million). The sample variances were $S_1^2 = .015$ and $S_2^2 = .008$. With $\alpha = .05$, we ask: Is there sufficient evidence to say that the population variances were different?

The F statistic from (12.1) gives

$$F = \frac{.015}{.008} = 1.875.$$

Consulting a table of percentage points for the F-distributions,[6] or using the cumulative distribution function for the F-distributions in `R`, we find a two-tail rejection region consisting of the union of the two intervals consisting of real values $F > 4.30$ together with $F < .2597$.[7] Since 1.875 is not in either interval, we do not have sufficient evidence to reject the null hypothesis here. $\triangle$

This indicates that we can proceed with the simplest form of the small sample test for difference of means. This is usually applied as follows. Since we want to proceed under the assumption that the variances of brain tissue DDT levels are equal in the two groups, the best way to estimate that variance is to *pool the two*

[6] We used the tables from Wackerly, Mendenhall, and Scheaffer, *Mathematical Statistics with Applications*, 6th edition here and in the later examples in this Chapter.

[7] Technical note: This is derived from the tabulated information as follows. With $11 - 1 = 10$ degrees of freedom in the numerator and $9 - 1 = 8$ degrees of freedom in the denominator,

$$F_{.025}(10, 8) = 4.30 \qquad \text{and} \qquad F_{.975}(10, 8) = \frac{1}{F_{.025}(8, 10)} = \frac{1}{3.85} \doteq .2597.$$

We are using a useful fact about the percentage points of the F-distributions here that is not difficult to derive with some algebra by manipulating inequalities:

$$F_{1-\alpha}(\nu_1, \nu_2) = \frac{1}{F_\alpha(\nu_2, \nu_1)}$$

(note that the numbers of degrees of freedom are reversed on the right hand side!)

groups (that is, to lump them together) and compute what is known as a pooled estimator for the common variance:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}.$$

Under the null hypothesis that the population means are *equal* it can be shown that the test statistic

$$T = \frac{\overline{Y_1} - \overline{Y_2}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

has a $t(n_1+n_2-2)$ distribution. Rejection regions for tests with any α are then derived as in our other examples.

EXAMPLE 12.10. We continue with the data from Example 12.9. The pooled estimator for the population variance is

$$S_p^2 = \frac{10\cdot(.015) + 8\cdot(.008)}{11+9-2} \doteq .012,$$

so $S_p \doteq .109$. This yields the test statistic

$$T = \frac{.039 - .024}{(.109)\sqrt{1/11+1/9}} \doteq .306.$$

Since the percentage point $t_{.025}(18) = 2.10092$, we see that $T = .306$ is very far from lying in the rejection region. This means that there is not sufficient evidence to reject the null hypothesis that the mean brain tissue DDT levels for the adults and the juveniles are actually the same. The fact that the adult level "looked" larger in these samples could be due to chance variation. △

12.6. Hypothesis Tests on Regression Coefficients

Although you might not have known it at the time, you have possibly seen p-values of hypothesis tests reported before. For instance, this happens when you use the regression functions from the Data Analysis package in Excel.

EXAMPLE 12.11. For example, suppose we enter the following data in Excel: in cells A1 - A8, $1, 2, 3, 4, 5, 6, 7, 8$ and in cells B1 - B8, $2.3, 3.6, 4.1, 4.3, 5.2, 5, 6.3, 7$. Highlight those cells and from the Data/Data Analysis menu, select Regression. The output generated (look on the tabs below for a new output sheet) includes something like this:

SUMMARY OUTPUT

Regression Statistics
Multiple R – 0.97434237
R Square – 0.949343055
Adjusted R Square – 0.94090023
Standard Error – 0.363787396
Observations – 8

⋮

	Coeff's	Std Error	t Stat	P − value	Lower 95%	Upper 95%
Int.	2.046428	0.283460	7.219439	0.00035	1.352824	2.740032
XVar.	0.595238	0.056133	10.60395	$4.142E-05$	0.457884	0.732592

What is going on here? Well, as we know, the R^2 value is an indication of how close the data was to lying on a single straight line (pretty close here!) In the block at the bottom, the Coefficients column shows the computed coefficients $\hat{b}$ and $\hat{m}$ in the regression line. This shows that the equation of the regression line is

$$y = \hat{b} + \hat{m}x = 2.046428 + 0.595238x.$$

The rest of that block reports the results of two statistical tests. Namely, the line for the $b =$ Intercept ("Int.") shows that in a test of $H_0 : b = 0$ versus $H_a : b \neq 0$, the p-value $p = .00035$ indicates quite strong evidence to reject H_0. Similarly, the last line gives the result of a similar test of $H_0 : m = 0$ versus $H_a : m \neq 0$. Here the results are even more striking, and we would conclude that there is very strong evidence for saying $m \neq 0$ from this data. This is often a key goal of a statistical regression analysis because in other words, it says that there is strong evidence y really does depend on x in a nontrivial way. The last two columns in the block give 95% confidence bounds for intervals containing m, b in an actual linear model describing the relation between y and x based on this data. $\triangle$

Since we have not discussed linear regression from quite this perspective before, in this section we want to describe some of the background behind the use of regression as a *statistical* technique and explain in general terms where this sort of information of p-values and confidence bounds on regression coefficients comes from. This is also the basis for some comments we made in Chapter 6 regarding the ways naive fitting of power law models via the log-log transform of the data and linear regression can lead to questionable results.

In our original presentation in Chapter 4, we described linear regression essentially simply as a data analysis tool, whereby we could find a linear model that fit a collection (x_i, y_i) of data points as well as possible. This of course ignores the possibility that the measured y_i for given values $x = x_i$ in a linear model can also be subject to chance variation following some probability distribution. If we take those sorts of random variations into account by incorporating a random variable *error term* ε in our linear models, we get something of the form:

$$\boxed{y = mx + b + \varepsilon.} \tag{12.2}$$

In this context, x is often called a *explanatory variable.* Then y is the *response variable* and y incorporates both deterministic and probabilistic contributions.

Assumptions 12.12. In this context, it is common to make the following simplifying assumptions regarding ε:

(1) For all x, ε has a known probability distribution with $E(\varepsilon) = 0$. As a result, for each x, y has a known probability distribution with expected value $E(y) = mx + b$. For instance, ε is often assumed to have a normal distribution with some variance σ^2 and mean $\mu = 0$.

(2) The variance σ^2 is *the same for all x*. This assumption has a rather forbidding name you might see if you look at other sources–*homoscedasticity*.[8]

To round out our discussion of issues related to fitting power law models from Chapter 6, understanding when a log-log transformed dataset satisfies these sorts of assumptions is not easy because there is little direct connection between these assumptions and properties of the original dataset that would be more natural to consider.

However, if we do know that Assumptions 12.12 hold, then the following result gives the basis for the kinds hypothesis tests on the regression coefficients that we saw in the output from Excel given above in Example 12.11. When we take a dataset (x_i, y_i) and compute the $\hat{m}$ and $\hat{b}$ coefficients in the equation of the regression line, we are "really" using the data to construct *estimators* for the coefficients m and b in the model (12.2). As always in Chapter 11, we can think of the estimators as random variables, and we have the following theoretical results.

Theorem 12.13. *Under Assumptions 12.12, and assuming ε has a normal distribution with constant variance σ^2, then*

(a) *$\hat{m}$ and $\hat{b}$ have normal distributions with expected values equal to the coefficients m and b in the model* (12.2).[9]

(b) *The variances are $V(\hat{m}) = c_{11}\sigma^2$ and $V(\hat{b}) = c_{00}\sigma^2$, where c_{11} and c_{00} are given by explicit formulas depending on the x_i from the data points.*

(c) *The quantity*

$$S^2 = \frac{1}{n-2} \cdot \sum_{i=1}^{n} (y_i - (\hat{m}x_i + \hat{b}))^2$$

gives a estimator for σ^2. The quantity $\frac{(n-2)S^2}{\sigma^2}$ has a χ^2 distribution with $n-2$ degrees of freedom. Moreover, S^2 is independent from both $\hat{m}$ and $\hat{b}$.

Then for tests of hypotheses about the slope of the regression line, with a value $m = m_0$ in the null hypothesis, one would use a test statistic

$$T = \frac{\hat{m} - m_0}{S\sqrt{c_{11}}}$$

which has a $t(n-2)$ distribution.[10] Similarly for tests about the y-intercept, with a value $b = b_0$ in the null hypothesis one would use

$$T = \frac{\hat{b} - b_0}{S\sqrt{c_{00}}}$$

which also has a $t(n-2)$ distribution. Upper-tail, lower-tail, and two-tail rejection regions based on given α would be found using the appropriate t-distribution.

It is the results of tests of this form that are reported by Excel and other statistical software. Since these computations are (almost) never done by hand in the real world today, we will not present a worked out example. See Example 12.11

[8]This comes from the Greek words for "same scattering," or "same variation."

[9]Technical note: This property is usually abbreviated by saying the estimators are *unbiased*.

[10]Technical note: This follows from the formulas given in footnote 11 from Chapter 11.

above and look at the Std Error and t Stat columns. Those were computed using the formulas we are describing.

12.7. χ^2 Tests for Model Fit and Independence

In this section we will present additional applications of the χ^2 distributions introduced in the final section of Chapter 11 to the analysis of categorical data and discrete probability models.

Tests of Model Fit. To motivate the sort of problem we will consider, consider the following situation.

EXAMPLE 12.14. Suppose we toss a standard six-sided die $n = 100$ times and record the following total number of times each of the possible rolls happened as in Table 1.

TABLE 1. Hypothetical data from $n = 100$ rolls of a die

1	2	3	4	5	6
21	15	13	16	12	23

We are dealing with *categorical* data here because each roll of the die gives us one of the 6 possible numbers $1, 2, 3, 4, 5, 6$. If we examine the data, since the numbers of 6's and 1's are seemingly much larger than the numbers of the other possible rolls, it is natural to consider the question: Do we have evidence here that the die is *"loaded"* (i.e. not a fair die)? $\triangle$

Another way to phrase this question is: Is this data consistent with the probability model of a fair die, where each of the six rolls is equally likely? The British statistician Karl Pearson (1857 - 1936) introduced the following method for addressing this sort of question. Phrasing this in the language of hypothesis testing, we want to develop a test with

- A null hypothesis, H_0: The data is consistent with a fair die.
- An alternative hypothesis, H_a: it is not.

More generally, we could consider similar questions related to the following sort of categorical data and a discrete probability model. Suppose we have a discrete random variable that can take some number k of possible different values, and we have probability mass function representing a model giving probabilities p_i, $i = 1, \ldots, k$ for each value. We sample some number n of values of the random variable, determine the *observed frequencies* O_i, and compare them to the and *expected frequencies* $E_i = p_i n$. We want to test:

- A null hypothesis, H_0: The data is consistent with the probability model, versus
- An alternative hypothesis, H_a: it is not.

Pearson introduced the following test statistic for measuring the deviation from an expected distribution.

$$\boxed{X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.} \tag{12.3}$$

He then established that the distribution of X^2 would be approximately a χ^2 distribution with $k-1$ degrees of freedom for k sufficiently large.[11] Even if n is not large, the percentage points of the appropriate χ^2 distribution can be used define rejection regions and possibly to estimate a p-value for an upper tail test. Let us see how this works using the data from Example 12.14 above.

Example 12.15. The data in Table 1 gives the O_i for $i = 1, \ldots, 6$. The E_i come from the probabilities for a fair die: $E_i = \frac{1}{6} \times 100 \doteq 16.7$ for $i = 1, \ldots, 6$. We compute

$$\begin{aligned} X^2 &= \frac{(21 - 100/6)^2}{100/6} + \frac{(15 - 100/6)^2}{100/6} + \frac{(13 - 100/6)^2}{100/6} \\ &\quad + \frac{(16 - 100/6)^2}{100/6} + \frac{(12 - 100/6)^2}{100/6} + \frac{(23 - 100/6)^2}{100/6} \\ &= \frac{146}{25} = 5.84 \end{aligned}$$

From a table of percentage points for χ^2 distributions, we see that with $6 - 1 = 5$ degrees of freedom, $\chi^2_{.05} = 11.0705$ and $\chi^2_{.1} = 9.23635$. Hence with $\alpha = .05$ or even $.1$, we do not have enough evidence to reject the null hypothesis that the observed counts are consistent with a fair die. The actual p-value of this test is $p \doteq .3221$,[12] which is definitely far too large to consider rejecting the null hypothesis. One way to interpret the p-value here is that if we repeated the experiment of rolling the die $n = 100$ times, then we would see unequal distributions of the six possible numbers *at least this extreme*[13] *only a bit less than* $1/3$ *of the time(!)* $\triangle$

Tests of Independence. Commonly-used hypothesis tests for independence are also based on the χ^2 distribution. We will begin by describing what we are looking for in this context. Saying two categorical random variables X, Y are independent should mean that for all possible pairs of values (x, y) in the joint sample space for X, Y, the events described by the equations $X = x$ and $Y = y$ are independent. In other words, from (10.4), we have

$$P(X = x \text{ and } Y = y) = P(X = x) \cdot P(Y = y). \tag{12.4}$$

---

[11]Technical note: This can be established using the Central Limit Theorem, but the details are far beyond the scope of this discussion.

[12]Technical note: This can be computed as

$$\int_{5.84}^{\infty} f(t)\ dt$$

using $f(t) =$ the $\chi^2(5)$ density and numerical integration software or by means of the functions in the `R` statistical package.

[13]in the sense that the X^2 statistic computed from the observed frequencies would be at least this large

Suppose we have gathered data about two such categorical variables, recording the total frequencies of all the possible pairs.

TABLE 2. Hair and Eye Color data

	Black	Brunette	Red	Blond	Total
Brown	68	119	26	7	220
Blue	20	84	17	94	215
Hazel	15	54	14	19	93
Green	5	29	14	16	64
Total	108	286	71	127	592

EXAMPLE 12.16. Table 2 shows data frequencies of combinations of hair color and eye color from a sample of $n = 592$ human subjects.[14] We might ask: Given this data, is it reasonable to conclude that the hair color (Black, Brunette, Red, Blond) and the eye color (Brown, Blue, Hazel, Green) are independent characteristics in humans? Or does the data favor the existence of some *association* between these traits? $\triangle$

In environmental science, similar questions come from studies of frequencies of different species of plants or animals in habitats with different characteristics and many other related questions.

The format of the data in Table 2 is an example of what is known as a *two-way contingency table*. The last column and the last row (labeled Total) give, respectively, the row sums and the column sums in the main, 4×4 section of the table at the upper left.[15] The sum of the row sums equals the sum of the column sums and that entry on the last row at the lower right gives the total sample size $n = 592$.

The basic idea of testing for independence with this sort of two-way categorical information is essentially to consider *all tables of the same size with the same row and column sums (and hence the same $n = 592$)*. It is easy to see that there are *an extremely large number of such tables*, since for instance we can (repeatedly) take any two rows and two columns in the 4×4 upper left section of the table and change the entries in the pattern $+1, -1$ in the first of the rows and $-1, +1$ in the second of the two rows, yielding new contingency tables with the same row and column sums.[16]

We will also need to consider such tables where the entries are not necessarily integers. Among such more general tables, we find exactly one corresponding to the independence condition. To describe this table, let us introduce the notation

[14] From R. Snee, "Graphical display of two-way contingency tables." *Amer. Statist.* **38** (1974), 9 - 12, quoted in: P. Diaconis and B. Efron, "Testing for Independence in a Two-way Table: New Interpretations of the Chi-Square Statistic," *Ann. Stat.* **13** (1985), 845-874.

[15] Technical note: These are often called the *empirical marginal distributions* of the two traits separately.

[16] Strictly speaking, we should also add the requirement that the subtractions leave entries that are still non-negative.

E_{ij} to represent the expected count under independence in row i and column j. Thinking about (12.4), we see that we want for a general n,

$$\frac{E_{ij}}{n} = \left(\frac{\text{row sum for row } i}{n}\right) \cdot \left(\frac{\text{colum sum for row } j}{n}\right),$$

so

$$E_{ij} = \frac{(\text{row sum for row } i) \cdot (\text{colum sum for row } j)}{n}.$$

TABLE 3. Expected Frequencies for Hair and Eye Color Under Independence

	Black	Brunette	Red	Blond	Total
Brown	$\frac{1485}{37}$	$\frac{7865}{74}$	$\frac{3905}{148}$	$\frac{6985}{148}$	220
Blue	$\frac{5805}{148}$	$\frac{30745}{296}$	$\frac{15265}{592}$	$\frac{27305}{592}$	215
Hazel	$\frac{2511}{148}$	$\frac{13299}{296}$	$\frac{6603}{592}$	$\frac{11811}{592}$	93
Green	$\frac{432}{37}$	$\frac{1144}{37}$	$\frac{284}{37}$	$\frac{508}{37}$	64
Total	108	286	71	127	592

EXAMPLE 12.17. Carrying out this process with the row and column sums from Table 2, you will check in Exercise 9 that the resulting independence table has the form given in Table 3. These could also be expressed using decimal approximations, of course, but these rational number values are *exact*. △

For the test of independence, we use a test statistic of the same form as in the test of fit. If there are r rows and c columns of observed frequencies data O_{ij} (excluding the row and column sums), then we consider

$$\boxed{X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}.} \tag{12.5}$$

In words, using the observed frequency O and the expected frequency E for each cell in the $r \times c$ main portion of the table (not including the row and column sums), we compute $\frac{(O-E)^2}{E}$ and sum them all.

It can be shown that for sufficiently large r, c, this X^2 statistic has a χ^2 distribution with $\nu = (r-1) \cdot (c-1)$ degrees of freedom. This gives the infrastructure needed for a test of

- A null hypothesis, H_0: The data is consistent with the independence model, versus
- An alternative hypothesis, H_a: it is not.

EXAMPLE 12.18. Using (12.5), you will check that $X^2 \doteq 138.29$ for the hair and eye color data. For $(4-1) \cdot (4-1) = 9$ degrees of freedom, from a table of percentage points of the χ^2-distributions, we find $\chi^2_{.005} \doteq 23.5893$. So we would reject the null hypothesis even with the very small $\alpha = .005$. In fact the actual p-value here is on the order of 2.3×10^{-25}, so there is very strong evidence to reject the null hypothesis that eye and hair color are independent. △

One might argue that in this case we are just verifying something that is obvious to any observant person–there is a perceptible association between certain hair colors and certain eye colors due to patterns in human genetics and the traits of different ethnic groups within the whole human population. In a sense, of course, that is true. However, our purpose here is to demonstrate a generally applicable statistical method that can be used to tease out associations in other cases that might not be as familiar and where our intuition and background knowledge do not necessarily provide any help.

12.8. Chapter Project

The Scientific Question. Let X be the lengths of male spiders of a particular large species and let Y be the lengths of female spiders of the same species (both in mm). A random sample of $n_X = 9$ X values were taken:

$$20.4, 21.7, 21.9, 21.4, 21.1, 23.6, 18.9, 22.6, 21.3$$

Similarly, a random sample of $n_Y = 13$ observations of Y were made:

$$20.5, 20.4, 20.3, 21.1, 21.2, 20.9, 21.0.21.3, 20.9, 20.0, 20.4, 20.8, 20.3$$

Is there a statistically demonstrable difference in the length distributions of two sexes, though? Is this species of spiders significantly sexually *dimorphic*[17] with respect to body size?

Answering the Question. If you think about discussion of small-sample confidence intervals for a difference of means from the text, you can see that there would be a corresponding hypothesis test. However, there were assumptions that we needed to check to apply the confidence interval formulas based on the t-distribution.

(1) What were they?
(2) There was one assumption about the overall distributions of the populations the sample values were coming from. Is there any reason to suspect that assumption is not satisfied for this data? Check using the `qqnorm` (normal probability plot) and `shapiro.test` (Shapiro-Wilk test) features of `R`. You can find all the information you need about format and syntax of these commands in many different places. Note: You should be able to interpret the output from `shapiro.test` by thinking about what we have said about hypothesis tests.
(3) There was also a second assumption about population variances. This will be tested using a hypothesis test based on an F-statistic as described in Example 12.9. Test the null hypothesis $H_0 : \sigma_X^2 = \sigma_Y^2$ against the alternative hypothesis $H_a : \sigma_X^2 \neq \sigma_Y^2$. You may select the significance level α. Interpret your results and give the attained significance level (p-value).
(4) Then, if there is *no demonstrable difference* in the variances, test for equality of the means using the t-test from the text. Use the pooled estimator S_p for the common variance, and test statistic

$$t = \frac{\overline{X} - \overline{Y}}{S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} .$$

[17] That is, occurring in two different forms.

For the alternative hypothesis $H_a : \mu_X \neq \mu_Y$ take the rejection region

$$RR = \{t \mid |t| \geq t_{\alpha/2}(n_X + n_Y - 2)\}.$$

(5) If there *is a demonstrable difference* in the variances, the basic t-test is not all that reliable in some cases. With small sample sizes, many experienced statisticians would use a different method due to Welch. Use the following test statistic:

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}},$$

but set up the rejection region for the test using a t-distribution with r degrees of freedom where

$$r = \left[\frac{(S_X^2/n_X + S_Y^2/n_Y)^2}{(S_X^2/n_X)^2/(n_X - 1) + (S_Y^2/n_Y)^2/(n_Y - 1)}\right] \tag{1}$$

($[z]$ = greatest integer less than or equal to z).[18] In this case, to test the null hypothesis $H_0 : \mu_X = \mu_Y$ against the alternative hypothesis $H_a : \mu_X \neq \mu_Y$, you would compute t as above and reject H_0 if $|t| \geq t_{\alpha/2}(r)$ where r is the number of degrees of freedom given in (1).

Note: You will need to select which of the two alternatives 4 and 5 you will use here, how you will set up the rejection region for your test, the test statistic, etc. based on the results of your test from part 1. Explain your choice. Clearly state the conclusion you draw from the test you carry out.

Assignment. Write up a report describing what tests you applied, what the results were, and your conclusion about the underlying scientific question about whether this species is *dimorphic* in body size.

Chapter Exercises

Directions for questions 1 - 4: For each question, first identify the relevant null and alternative hypotheses. Then carry out the appropriate hypothesis test, estimate the p-value if possible with the information you have, and answer the other questions.

(1) CO_2 emissions per Megawatt-hour of electric power produced were measured for $n = 100$ coal-fired power plants. The average CO_2 in lb/Mwh is $\overline{Y} = 2223.1$ with $SD = 211.3$ lb/Mwh.

   (a) If this were a random sample, would there be evidence at the $\alpha = .05$ level to say that the average coal-fired power plant emits an amount different from 2200 lb/Mwh?

   (b) The data used in this problem actually comes from the 100 largest coal-fired electric power plants in the U.S. (this is true). Is this a random sample? Does the calculation in part a) really make sense in this setting? Discuss.

[18] Technical note: The difference here is that the two samples are *not* being lumped together to compute the pooled estimator for the variance, but the number of degrees of freedom in the t-distribution is being adjusted to account for the two different estimates S_X^2 and S_Y^2.

(2) A simple random sample of 2000 Germans in 2001 showed that 840 thought that all electricity should be generated by "green" energy sources such as wind and solar power. Is there evidence at the $\alpha = .05$ level to say that at least 40% of all Germans thought the same about sources of electric power?

(3) In a study of $n = 17$ sea star arm lengths, the average was $\overline{Y} = 6.8$ cm with an SD $= .5$ cm. Is there evidence at the $\alpha = .05$ level to say that the population arm length is different from 7 cm? Estimate the p-value of the test using the information in Figure 11.2.

(4) A medical study compared the resting pulse rates of random, independent samples of 100 smokers and 100 nonsmokers. The smokers had an average pulse rate of $\overline{Y} = 86$ beats per minute with $SD_1 = 5.4$, while the nonsmokers had an average pulse rate of $\overline{X} = 80$ and $SD_2 = 4.9$. Is there evidence at the $\alpha = .05$ level to say that nonsmokers have a lower average pulse rate?

(5) Criticize the following statement. Is it true? Is it false? Is it a reasonable way to think about what the p value means? "*The p-value of a statistical test of hypotheses is the probability that the null hypothesis is true.*"

(6) (Hypothetical data) A new diet regimen is being tested against an old diet for farmed tilapia.[19] To compare the eventual body weights of the fish raised under the old and new diets, random samples with $n = 30$ are selected from fish raised with each diet. Each fish is weighed after reaching 240 days in age (full growth). Let X be the weight of a fish raised on the new diet, and Y be weight of a fish raised on the old diet, both in units of grams. The goal is to determine whether the new diet increases the adult body weight of the fish. The data collected was as follows:

	1250	1210	990	1310	1320	1200	1290	1360	1200	1150
X :	1120	1360	1310	1110	1320	980	950	1430	1100	1080
	960	1050	1310	1240	1420	1170	1470	1060	1230	1300

and

	1180	1360	1310	1190	920	1060	1440	1010	1000	950
Y :	1310	980	1310	1030	960	800	1280	1080	900	1030
	930	1050	1010	1310	940	860	1450	1070	840	1100

(a) Construct parallel "box-and-whisker plots" for these data sets as in Chapter 9 and make an informal conjecture about whether or not the new process (the X data) has increased the fish body weight, compared with the Y data.

b) Describe an appropriate test of the null hypothesis $H_0 : \mu_X = \mu_Y$ versus $H_a : \mu_X > \mu_Y$. Say what your assumptions about the data are, what your test statistic is, what the rejection region will be, and so forth.

c) Carry out your test at the $\alpha = .01$ level of significance. Give a clear and concise statement of the conclusion you draw from your test.

[19]Tilapia are an important aquaculture species. They are fast growing, lean and short-lived, so they do not accumulate mercury and other toxins to the same level that other food fish do. They are a good protein source, but their flesh generally does not contain as much of the beneficial omega-3 fatty acids as other food fish species such as salmon.

d) What is the attained significance level of your test (the p-value)? What does this say?

(7) A researcher found that in two independent random samples, 107 out of 200 apple seeds germinated when they were planted in soil maintained at $5°C$, while 123 out of 200 seeds germinated when they were planted in soil maintained at $15°C$.

(a) Construct a 95% confidence interval for the difference in the proportions of seeds that germinate at the two temperatures using our discussion of confidence intervals from Chapter 11. State any assumptions you are making.

(b) Do the results of this study suggest that a greater proportion of apple seeds will germinate at the higher temperature in general? Construct an appropriate test of hypotheses, carry it out, and interpret your results. How does your answer relate to what you found in part (a)?

(8) Consider the following measurements of the heat-producing capacity of the natural gas produced by two fields of gas wells (in calories per cubic meter):

Field 1 :	8.26	8.13	8.35	8.07	8.34	
Field 2 :	7.95	7.89	7.90	8.14	7.92	7.84

Let μ_i $(i = 1, 2)$ be the population mean heat-producing capacity of the natural gas from field i.

(a) What assumptions do you need to make in order to use the appropriate test of $H_0 : \mu_1 = \mu_2$ versus $H_a : \mu_1 \neq \mu_2$? How could you determine whether it is reasonable to assume those assumptions are satisfied?
(b) Set up and carry out appropriate tests to determine whether the assumptions from part (a) are satisfied.
(c) Carry out the test on the population mean heat-producing capacities and state your conclusion clearly and succinctly.

(9) Verify the form of the expected frequencies (E_{ij}) in Table 3.

(10) Check the value $X^2 \doteq 138.29$ obtained from Tables 2 and 3. Which of the terms in the sum from (12.5) are especially large in this case? What do these large terms mean in real world terms?

(11) Suppose we have an $r \times c = 2 \times 2$ two-way contingency table as in Table 4 with all $a, b, c, d > 1$. The row and column headings (i), (ii), (I), (II) just indicate that there are two possible values for each of the categorical variables.

TABLE 4. A general 2×2 two-way contingency table

	(i)	(ii)	Total
(I)	a	b	$a + b$
(II)	c	d	$c + d$
Total	$a + c$	$b + d$	$a + b + c + d$

Show that the X^2 statistic in this case is computed by the shortcut formula:

$$X^2 = \frac{n(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$

Hint: This is just a big algebra exercise(!) Unfortunately, for larger tables, this kind of approach quickly becomes totally unwieldy.

(12) A New York Times/CBS poll conducted between April 5 and April 12, 2010 included the question "Do you approve or disapprove of the way your Representative in Congress is handling his or her job?" The poll was carried out in two stages. First a large national simple random sample was asked the question. The results were reported in three categories: 46% of respondents said they approved, 36% said they disapproved, and 18% said they had no opinion. Then the same question was asked of a separate sample of $n = 881$ people, all of whom had identified themselves as supporters of the "Tea Party" movement. The responses broke down like this in the same three categories:

	Approve	Disapprove	No opinion
Tea Party Sample	352	432	97

(Data is adapted from the New York Times web site.) It seems from the numbers that "Tea Party" supporters may differ from the general population when it comes to their opinions concerning their Congressional Representatives. But is this a real difference, or could it just be a product of chance variation in the sampling process? Test using an appropriate χ^2 test of model fit.

(13) In a study of bird populations in different habitats in California, Latta et al.[20] collected samples in "remnant" habitats (that, is areas along rivers with mostly native vegetation) and "restored" habitats (environmentally damaged areas where native vegetation was replanted by humans). They observed the data of numbers of birds of different species in each type of habitat (the "Uncommon" category is a catch-all for less-common species) given in Table 5.

Do these data support the conclusion that there is an association between the frequencies of different bird species and the type of habitat, or are those two variables independent?

[20] Latta, S.C., C.A. Howell, M.D. Dettling, and R.L. Cormier. "Use of data on avian demographics and site persistence during overwintering to assess quality of restored riparian habitat," Conservation Biology **26** (2012), 482-492.

TABLE 5. Bird counts in remnant and restored habitats

Species	Remnant	Restored
Ruby – crowned kinglet	677	198
White – crowned sparrow	408	260
Lincoln′s sparrow	270	187
Golden – crowned sparrow	300	89
Bushtit	198	91
Song sparrow	150	50
Spotted towhee	137	32
Bewick′s wren	106	48
Hermit thrush	119	24
Dark – eyed junco	34	39
Lesser goldfinch	57	15
Uncommon	457	125

Index