1

Philosophical and Historical Foundations

The most beautiful thing we can experience is the mysterious. It is the source of all true art and all science. He to whom this emotion is a stranger, who can no longer pause to wonder and stand rapt in awe, is as good as dead: his eyes are closed.

Albert Einstein (1931)

1.1 Introduction

Before beginning a book on statistics for biologists, it will be helpful to take a general view of the concepts that underlie science. It seems reasonable that biologists should be able to understand, to some degree, the general characteristics and goals of science. Further, as scientists using statistics, we should be able to understand the way these goals are addressed by statistical tools. We should bear these ideas in mind as we design our experiments, collect our data, and conduct our analyses.

1.2 Nature of Science

Based on its etymology, *science* (from the Latin *scire* meaning "to know") represents a way of "knowing" about our universe. This conceptualization, however, is still vague. What do we mean by "knowing," and what do we mean by "a way of knowing"; is there more than one way of knowing something? Such philosophical questions go beyond the scope of a book on statistics. Nonetheless, statistics can be considered a practical response to these ideas. For instance, probability, the language of statistics, can be used to quantify both the "believability" of a scientific proposition, and the amount of knowledge/evidence we have concerning a phenomenon.

Because attempts at an all-encompassing definition of science gives rise to still other questions, many authors emphasize only the methodological aspects of science. For example, standard dictionary definitions stress procedures for the accumulation of knowledge. These include "a systematic and formulated knowledge,"* "systematized knowledge derived from observation, study, and so on,"† and "knowledge covering general truths of the operation of general laws, concerned with the physical world."‡ Thus, science is concerned with knowledge: a special kind of technical knowledge that may be

* *The Oxford Pocket American Dictionary of Current English.*
† *Webster's New World Dictionary.*
‡ *Merriam Webster's New Collegiate Dictionary*, 11th edition.

intrinsically trustworthy because its methods emphasize precision and objectivity, and useful, because it describes the everyday phenomena of the universe.

Out of necessity, this chapter also explores science methodologically. I briefly address modes of scientific description, the structure and history of the scientific method, the relationship of science to logic, and the correspondence of statistics to the goals and characteristics of empirical science. This approach, however, ignores several fundamental issues. For instance, it does not explain why humans are driven to *do* science. To this end, it has been suggested that science satisfies a basic human need (cf. Tauber 2009). Through its program of discovery, science may provide meaning—a personal context in a societal setting, the biosphere, and the universe.

1.3 Scientific Principles

Science can be loosely subset into the *empirical sciences,* which are concerned with suppositions that are empirically supported (e.g., biology) and the *nonempirical sciences* whose propositions are supported without empirical evidence (e.g., logic and pure mathematics).[*] More than any other enterprise, the field of statistics bridges these branches by quantifying the character of empirical data with mathematical tools.

Scientists have developed specific procedures for observing and considering phenomena, along with unique methods for description (i.e., scientific pictures[†]) that distill technical knowledge into practical forms. Three principles guide these processes: objectivity, realism, and communalism.

1.3.1 Objectivity

The process of science emphasizes *objectivity.* That is, an attempt is made to study and describe phenomena in a way that does not depend on the investigator (Gaukroger 2001). Objectivity has been a primary motivator for the development of the scientific method (Section 1.4) and the statistical tools described in this book, since both seek to provide rigorous and impartial methods for describing data and making inferences (Hempel 1966, p. 11). To facilitate objectivity, a number of procedural steps are often taken by scientists. For example, methods and tools that are not biased (predisposed to certain answers at the expense of other equal or more valid answers) should be used for measurements and descriptions. With respect to sampling and experimental design, randomization can be used to select samples representing a phenomenon, and to assign these samples to experimental treatments (see Chapter 7). These steps will tend to decrease investigator bias, and balance out the confounding effect of unmeasured variables.

1.3.2 Realism

Scientists also strive for *realism* (recognizable depictions) in their research. The doctrine of *scientific realism* holds that (1) products of science exist independently of phenomena

[*] The empirical sciences can be further subset into the *natural sciences,* including biology, chemistry and physics, and the *social sciences* encompassing anthropocentric topics like sociology, anthropology, and history.

[†] For discussion and comparisons of art and science, see Stent (1972), Hofstadter (1979), and Cech (1991).

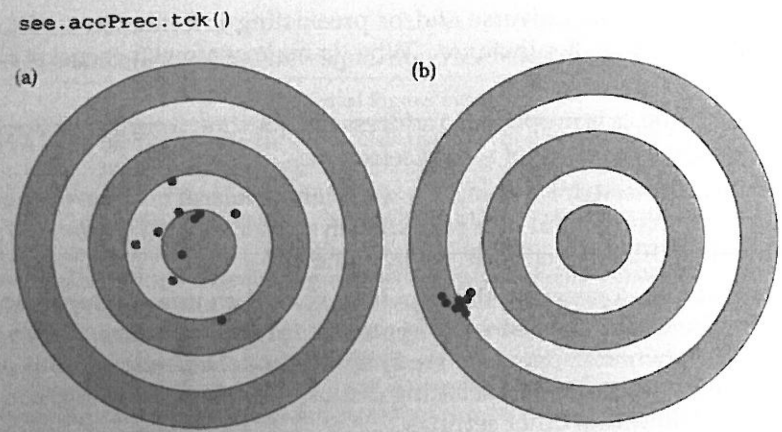

FIGURE 1.1
The concepts of accuracy and precision can be considered interactively using the **R** function `see.accPrec.tck` in asbio: (a) high accuracy but low precision; (b) high precision but low accuracy.

being studied, and (2) scientific theories that have been repeatedly corroborated represent truth-directed descriptions of reality (Boyd 1983, Popper 1983, Psilos 1999). Thus, one measure of realism in science is repeatability (i.e., verifiability, *sensu* Hempel 1990). To facilitate repeatability, so that others may substantiate or disprove their results, scientists (1) carefully document their methods of research (including their methods of analysis), and (2) use methods of measurement that are both *accurate* (i.e., give readings close to a "true" reference value) and *precise* (i.e., give very similar measures if repeated for the same object) (see Figure 1.1).

1.3.3 Communalism

The growth of science depends on the refinement, corroboration, and refutation of scientific claims. Because these actions require a community of participants, scientific knowledge can be considered a *communal* process (cf. Shapin and Schaffer 1985). Each scientist is a member of this community, and a contributor to a picture of the universe that has morphed and grown since the dawn of knowledge.

1.4 Scientific Method

While it is helpful to describe general characteristics of scientific description (objectivity, realism, communalism), these are insufficient as a guide for conducting scientific research. Such a framework, however, is provided by the *scientific method*. The process is often presented in general biology texts* as four progressive steps:

* For example, Freeman (2010) and Reece et al. (2011).

1. Observations about the universe and/or preexisting information lead to the formulation of a question. For instance, "Why do male peafowl (peacocks) have elaborate tails?"
2. A testable hypothesis is proposed to address the question from step 1. For instance, "Peacock tails are the result of sexual selection."
3. The hypothesis is tested. For instance, we might measure tail size for males with successful and unsuccessful courtship displays. If males with larger tails are more successful, then this test supports the hypothesis.
4. A decision is made concerning the hypothesis based on information from the test. Regardless of whether the test supports or refutes the hypothesis, more information about the phenomenon under study is acquired. Depending on the decision, additional steps can be taken, including disposal, revision, retesting, and application of the hypothesis to other settings.

This general format has existed for thousands of years. However, perspectives on underlying issues have changed dramatically over this period. These include answers to important questions like "what constitutes evidence?"; "what is the role of experimentation?"; "how do we quantify evidence given that experimental outcomes vary?"; and "how do we weigh evidence from past observations with respect to results from current experiments?"

1.4.1 A Terse History

Aristotle (384–322 BC) is often referred to as the father of the modern scientific method because of his support for a version of *empiricism*, the view that knowledge comes from sensory experience (Pedersen 1993). Aristotle supported the idea that *premises* (propositions prompting a specific conclusion) for "true scientific enquiry" could come from observation. He insisted, however, that genuine scientific knowledge could only come from the application of pure reason, undiluted by observation. Aristotle's empiricism, while limited, was a compromise with the views of his teacher Plato (428–348 BC) who thought that empiricism was dangerous, and that knowledge could only come from rational thought and logic.

While not often considered in histories of science, Aristotle and Plato were preceded by Grecian philosophers such as Thales of Miletus (624–c. 546 BC) and Empodocles of Agrigentum (c. 490–430 BC) who espoused both observation and experimentation. These approaches were, in turn, greatly predated by Egyptian records from at least the 17th century BC describing the basic tenets of modern science (Achinstein 2004).[*] Still earlier, Babylonian and Mesopotamian methods of astronomy from the dawn of written history (ca. 3100 BC) provide a basis for "all subsequent varieties of scientific astronomy … if not indeed all subsequent endeavors in the exact sciences" (Aaboe 1974).

[*] Histories of science are often Western-centered. For instance, the Persian *Avicenna* (980–1037 AD), a Muslim physician, scientist, and philosopher outlined a sophisticated and modern version of empiricism that emphasized controlled experimentation as opposed to mere mensurative observation. His contributions, which predate the Renaissance, have been largely overlooked in Western histories of science.

TABLE 1.1
Major Modern (Renaissance–Present) Developments in the Scientific Method

Historical Figure/Event
Galileo Galilei (1564–1642) used work from earlier scientists and his own experiments to support the hypothesis of a Copernican heliocentric solar system. Galileo's empiricism was unique for his time. In the late Renaissance, scientific knowledge was mainly the product of axiomatic deduction from fundamental principles or common notions. As a result, Galileo concealed the role of both experimentation and induction in his conclusions.
Francis Bacon (1561–1626), a contemporary of Galileo, helped to formalize the scientific method and was a strong advocate of induction and empiricism in scientific research. Bacon's penchant for inductive reasoning was demonstrated by the importance he placed on written histories of natural and experimental phenomena. These were to comprise part of his so-called *great instauration* or great restoration of science.
Isaac Newton (1642–1727), originator of many profound ideas in math and physics, demonstrated that mathematics and experimentation could be combined in scientific research. Newton's *Opticks* listed 35 experiments, including a famous test demonstrating that color is the result of objects interacting with light rather than color being generated by the objects themselves. Newton's views of the scientific method were important. He favored an approach in which conclusions were reasoned inductively and deductively from experimental evidence without the use of *a priori* hypotheses. Newton was also opposed to the use of "probable opinion" in research because the aim of science is certain knowledge.
A number of Newton's contemporaries did not share his negative opinion of "probable opinion." For instance, *Christian Huygens* (1629–1695) in his *Treatise on Light* recognized that his "demonstrations did not produce as great a certitude as those of geometry" because premises used by geometers are fixed and based on deductive "incontestable principles." Nonetheless, Huygens claimed a high-degree certainty with his results because his conclusions were supported by a "great number" of his observations.
A large number of mathematicians and scientists, including *Pierre de Fermat* (1601–1665), *Abraham de Moivre* (1667–1754), *Blaise Pascal* (1623–1662), *Gottfried Wilhelm Leibniz* (1646–1716), *Simon Pierre de Laplace* (1749–1827), *Jakob Bernoulli* (1654–1705), *Siméon-Denis Poisson* (1781–1840), *Daniel Bernoulli* (1700–1782) helped develop a mathematical method for quantifying the practical certainty of hypotheses. This approach, called probability, allowed one to view evidence for or against hypothesis as a continuum.[a]
John Maynard Keynes (1883–1946) defined new roles for probability in science by drawing on the mathematical ideas of *Thomas Bayes* (1702–1761). In particular, Keynes claimed that the probability for a hypothesis could be viewed as an investigator's degrees of belief in that hypothesis. Keynes also suggested that prior probabilities could be assigned to a hypothesis representing background knowledge. In this *Bayesian* approach, prior probability could be augmented with current evidence in an inductive manner.
Hans Reichenbach (1891–1953) offered a new perspective on probability. Specifically, he argued that the probability of an event comprised its long-run relative frequency. The *frequentist* view fits well with the use of probability in most modern statistical procedures.
Karl Popper (1902–1994) favored the use of deduction in hypothesis evaluation in his *hypothetico-deductive* approach (Section 1.6.4). Popper's position for deduction (and against induction) was due to *David Hume* (1711–1776) who insisted that the logic of induction is not justifiable because it relies on the "uniformity of nature" (Hume 1740, *qtd in* Selby-Bigge 1996). That is, what happened in the past cannot be assumed to be a perfect predictor for what will happen in the future. However, Popper's view of hypothesis corroboration (repeated nonfalsification of hypotheses) has caused many philosophers of science to claim that Popper's approach was "inductive in the relative sense" (Salmon 1967).

[a] For an overview of the history of statistics and probability, see Owen (1974).

A few figures important to the modern (Renaissance to present) development of the scientific method are described in Table 1.1.* The table reveals three trends that are explored in greater detail throughout the remainder of the chapter. First, there has been an increased reliance on experimentation as a method for testing hypotheses. Second, a greater emphasis has been placed on a form of logic called induction. Third, since its invention in the 17th

* A detailed account of the development of the modern scientific method is beyond the scope of this book. Interested readers are directed to Lloyd (1979), Pedersen (1993), Gower (1997), and Achinstein (2004).

century, probability has been used increasingly, and in increasingly sophisticated ways, to quantify evidence concerning scientific claims.

1.4.1.1 Experimentation

To be taken seriously by his contemporaries, Galileo Galilei (1564–1642) concealed the use of experimentation in his work concerning gravity (Table 1.1). Instead, he attempted to emulate the axiomatic approach of Euclid (fl. 300 BC) in which self-evident truths were deduced from "common notions" (Gower 1997). For instance, instead of summarizing data from his experiments involving balls of different weights rolling down an inclined plane, Galileo presented a thought problem in which two objects, one heavy and one light, were dropped simultaneously from the top of a tower (Galilei 1638). Hypothetically, if the heavier object reaches the ground first, because its greater weight results in greater free falling velocity, then tying the lighter object to the heavier object should slow the heavier object down (and accelerate the lighter object). However, if weight determines velocity, as was first assumed, then the conjoined objects should travel faster than the heavier object alone. Because these arguments are in contradiction, Galileo argued that weight does not influence acceleration due to gravity.

The resistance to experimentation, demonstrated by Galileo, was due to the enduring views of Aristotle and Plato, and the fact that earlier methods of measuring phenomena were imprecise and unrepeatable (Gower 1997). As measurement tools and techniques were improved, and the practical benefits of experiments were demonstrated—for instance, through tests that resulted in effective treatments for disease (Hempel 1966, p. 3)—trust in experimentation increased.

1.4.1.2 Induction

With *inductive reasoning*, we draw conclusions concerning a phenomenon based on accumulated evidence, while acknowledging that the evidence is incomplete (induction is fully explained in Section 1.6.1). The resulting conclusions may be considered "probably correct," but not certain, because they are based on imperfect and incomplete observation. This restriction led many early philosophers to disparage the use of induction in science. However, trust in induction has grown with its demonstrated usefulness, and with improvements in methods for recording, storing, and analyzing data. Modern critics of induction, particularly Karl Popper (1902–1994), have themselves been criticized for refusing to acknowledge the merits of induction, while tacitly embracing its tenets.

1.4.1.3 Probability

Two major conceptions of probability have arisen over time. The first, formalized by Hans Riechenbach (1851–1953), views probability as a limiting frequency. That is, the probability of any event is its relative frequency over an infinite number of trials. This view is consonant with the most commonly used approaches in statistics, including point and interval estimation of parameters (Chapters 4 and 5), and null hypothesis testing (Chapters 6 through 11). The second, attributed to Thomas Bayes (1702–1761), views the probability of an event as the degrees of belief in the truth of that event. Mathematical ideas that underlie this perspective and an increasing trust in induction have resulted in useful methods that create probabilistic statements by synthesizing prior empirical knowledge and current knowledge. The importance of probability to science is revisited in Section 1.8. Probability itself is formally defined and described in Chapter 2.

1.5 Scientific Hypotheses

A *hypothesis* is a statement describing the universe that serves as a starting point for a scientific investigation. All branches of science, including biology, are built on systems of hypotheses. We may be most familiar with hypotheses of causality (e.g., *this* causes *that*) or hypotheses that are comparative (e.g., *this* and *that* are different, or *this* is bigger or smaller than *that*). Hypotheses, however, allow other considerations of phenomena. For instance, they are often used to address association (e.g., *this* increases with *that*), or description, (e.g., *this* is a true characteristic of *that*), or to provide a means for answering, which, in a series of models, is the best at explaining a phenomenon (e.g., *this* is the best model we have for *that*).[*]

Scientific and nonscientific hypotheses can be distinguished using two criteria. First, a scientific hypothesis will be *testable*. In using the term testable, I mean that an experiment can be designed that provides information concerning the validity of the hypothesis. Of course, the scientific method as described in Section 1.4 explicitly requires testability. Second, many scientists have argued that scientific hypotheses should be falsifiable. *Falsifiability*, popularized by Karl Popper, requires a conceptual outcome that would cause a scientific hypothesis to be rejected. That is, "A theory is potentially a scientific theory if and only if there are possible observations that would falsify (refute) it" (Boyd et al. 1991).

1.5.1 Falsifiability

Popper himself did not advocate discarding a strong hypothesis as soon as negating evidence was encountered, although he has been linked to this idea (cf. Okasha 2002).[†,‡] To quote Popper (1959, p. 32): "once a hypothesis has been proposed and tested, and has proved its mettle, it may not be allowed to drop out without a 'good reason.'" That is, extraordinary falsifications require extraordinary proof (Popper 1959, p. 266).[§] Popper did, however, favor severe testing for scientific hypotheses, and argued that no hypothesis should be immune from scrutiny. That is, "...he who decides one day that scientific statements do not call for any further test, and that they can be regarded as finally verified, retires from the game" (Popper 1959, p. 32).

---

[*] Science (and statistics) is not solely concerned with hypothesis testing. For example, each year, a large number of descriptive scientific papers are published that do not explicitly test hypotheses (although they may serve as material for hypothesis generation). Still, other scientific studies are concerned with the invention or improvement of tools or methods, for example, measuring instruments and technological advances (although they may be later used to quantify evidence concerning scientific hypotheses).

[†] See Sober (1999, pp. 46–49) for criticisms and comments on falsification in a biological context.

[‡] It can be demonstrated that important advances in science have occurred when the evidence against a hypothesis *did not* lead to its immediate falsification. For instance, astronomers in the 19th century found that the orbit of Uranus was poorly described by Newtonian physics. Instead of rejecting Newton's ideas, however, John Couch Adams and Urbain Leverrier independently predicted that an undiscovered planet was the cause of the orbital discrepancies. Neptune was soon discovered in the exact region forecast by the scientists. On the other hand, consider two other scientific breakthroughs, the theory of relativity and quantum mechanics. These developments demonstrate that Newton's physics gives incorrect results when applied to massively large objects, objects moving at high velocities, or subatomic particles. Blind adherence to Newtonian theories (nonfalsifiability) could have prevented these important discoveries.

[§] Strong scientific hypotheses that withstand repeated attempts at falsification are often termed *theories* (Ayala et al. 2008).

Two notable philosophers have presented conceptions of science counter to Popper's severe falsification. They were Imre Lakatos (1922–1974) and Thomas Kuhn (1922–1996). Lakatos (1978) argued that severe falsification was at odds with the possibility of scientific progress. As an alternative, he suggested that scientific theories constituting a "research programme" be classified into two groups. The first would consist of a core set of theories that would be rarely challenged, while the second would be comprised of a protective "auxiliary layer" of theories that could be frequently challenged, falsified, and replaced (Mayo 1996, Quinn and Keough 2002). Kuhn (1963) showed that, in general, falsified hypotheses do not result in the rejection of an overall scientific paradigm, but instead serve to augment and develop it. This state of "normal science" is only occasionally disrupted by a "scientific revolution" in which a paradigm is strongly challenged and overthrown. Kuhn has been described as a critic of scientific rationality (Laudan 1977). However, Kuhn (1970) emphasized that his goal was not to disparage science, but to describe how it worked. That is, science does not progress linearly, but often stumbles and lurches eccentrically (Okasha 2002).

The Popperian, Lakatosian, and Kuhnian conceptions of science each have strengths. Popper presented a clear method for research—propose a hypothesis, and then do everything possible to reject it. If the hypothesis withstands severe testing, then it can be considered a provisional characterization of the universe. Kuhn did not propose a method for scientific research as such, but described how science can proceed nonlinearly (Kuhn 1970). Lakatos proposed taking core theories off the table, something that Popper would have disagreed with, but his views of hypothesis testing were more sophisticated than Popper's (Hillborn and Mangel 1997). In particular, Lakatos proposed that multiple competing hypotheses will always be possible for explaining a phenomenon. As a result, an extant hypothesis is never rejected, but only replaced by a better "research programme." This is in contrast to the Popperian severe tests of individual hypotheses.

1.6 Logic

The approach used by scientists to formulate hypotheses and understand the world is rooted in the precepts of logic. There are two modes of logical reasoning: induction and deduction. Both can be summarized using verbal arguments consisting of *premises* and *conclusions. Logic* itself can be defined in the context of a these arguments. Specifically, an inductive or deductive statement is said to be valid or *logically correct* if the premises of the argument support the conclusion (Salmon 1963). Conversely, if the premises do not support the conclusion, then the argument is *fallacious*.

1.6.1 Induction

Formalization of the scientific method by Francis Bacon (Section 1.4.1) and John Stuart Mill (1806–1873) emphasized the importance of *induction*. In this framework, empirical or conceptual premises (e.g., observations, existing theory, and personal insights) lead to a conclusion that is *probably true*, if the premises are true (Russell 1912, Keynes 1921). Here is an example of inductive reasoning:

No bacterial cells that have ever been observed have nuclei.	} Premise
Bacteria do not have nuclei.	} Conclusion

Because of their reliance on uncertain or incomplete premises, inductive arguments have three distinguishing characteristics. First, the conclusion will contain information not present, even implicitly, in the premises. Second, statements in the premises will be more specific compared to more general statements in the conclusion. Third, the conclusion may be false even if the premises are true.

1.6.2 Deduction

Other philosophers of science, notably René Descartes (1596–1650), emphasized the importance of *deduction*. Here, one states a conclusion already contained in the premises (cf. Hempel 1966). An example of deductive reasoning is shown here.

Bacterial cells do not have nuclei.	} Premise 1
Escherichia coli (*E. coli*) are bacteria.	} Premise 2
E. coli do not have nuclei.	} Conclusion

Deduction has two distinguishing characteristics. First, deductive arguments lead from general premises to a more specific conclusion. For example, in the argument above, the first premise is a statement concerning all bacteria, while the conclusion concerns only one bacterial species. Second, if deductive premises are true, then the conclusions from a logically correct deductive argument *must be true*.

It is clear that mathematical arguments are based on deduction. In a mathematical proof, axioms (which are assumed to be true) are used to derive a conclusion (e.g., a theorem), which, by definition, must also be true.*

1.6.3 Induction versus Deduction

Induction has been criticized as irrational because "believing" in its conclusions means "believing" that the universe never changes. That is, what we learned yesterday is completely applicable today, tomorrow, and so on. Hume (1740) formalized this issue in his *problem of induction*. Many authors, however, have argued that Hume's problem is irrelevant since induction constitutes a fundamental component of human cognition, and that not using it would be crippling (Strawson 1952). Indeed, induction has been increasingly trusted as science has developed.

---

* It is interesting to note that the validity of an argument does not require that that premises are true. For instance, the following represents an argument that is scientifically incorrect (from the perspective of the 21st century), but logically correct:

Flat objects have edges. } Premise 1
The earth is flat. } Premise 2
Traveling a long distance in a straight line will cause one to fall off the edge of the earth. } Conclusion

A fallacious argument can also have both "true" premises and conclusions. Consider the following deductive example:

All mammals are air breathing.	} Premise 1
All cats are air breathing.	} Premise 2
All cats are mammals.	} Conclusion

While the premises and conclusion are both true, the general form of the premises do not support the conclusion. For instance, lizards are air breathing, but are not mammals.

It seems obvious that modern science values and uses both induction and deduction (Salmon 1967, Mentis 1988, Chalmers 1999, Okasha 2002, many others). Deduction is useful for many applications, including mathematics, and null hypothesis testing (see Section 1.6.4). Deductive premises, however, may be the product of induction. Furthermore, the creation of hypotheses from supporting observational evidence is by definition inductive. This association is vital to the establishment of scientific theories and laws (Ayala et al. 2008, p. 11).

EXAMPLE 1.1

Consider a wildlife biologist in Yellowstone National Park. The biologist observes that narrowleaf willow (*Salix angustifolia*) seedling recruitment always seems to decrease as elk (*Cervus canadensis*) populations increase in size. Based on this pattern, she inductively hypothesizes that elk negatively affect willow survival. She further reasons that willow seedlings protected by elk exclosures should have higher recruitment than adjacent groves (controls), which are freely accessible to elk. She gathers appropriate data and tests this hypothesis using mathematical/statistical methods. This would be a test of a deductive argument based on premises from inductive reasoning.

1.6.4 Logic and Null Hypothesis Testing

It is possible to make connections between the precepts of logic and the methods of statistical hypothesis testing described later in this book. In this section, we will examine the logical foundations of an approach called *null hypothesis testing*, that is, testing a hypothesis of "no effect." The null hypothesis testing paradigm is fully explained in Chapter 6.

1.6.4.1 Modus Tollens

Deduction fits well with severe falsification because rejection of a hypothesis based on data is always possible, but empirical confirmation is impossible (Popper 1959). This is because the premises cannot constitute all possible data. To illustrate this, let H be a hypothesis, and let I be an entity that allows evaluation of H. Now, consider the following arguments:

If H is true, then so is I	} Premise 1	
But available evidence shows that I is not true	} Premise 2	(1.1)
H is not true	} Conclusion	

If H is true, then so is I	} Premise 1	
Available evidence shows that I is true	} Premise 2	(1.2)
H is true	} Conclusion	

In Argument 1.1, we reject hypothesis H using a *logical* form of deduction called *modus tollens*. The statement is deductive because if the premises are true then the conclusion must be true as well. This form of argument is also called *denying the consequent* because the consequence of H is denied, resulting in the refutation of H.

The argument in Argument 1.2 is a form of deduction called *affirming the consequent*.* In this case, the conclusion may be false even if the premises are true. Thus, the argument is fallacious. The first premise indicates that I is dependent on H, not the converse. Thus, the truth of I (suggested in the second premise) may not signify the truth of H. In addition, the second premise is inconclusive because it consists of incomplete evidence (Hume 1740). Because at least some information concerning H is unknown, we cannot prove that the hypothesis H is true. At best we can say that we have failed to reject the hypothesis.

Thus, we can only deductively reject or fail to reject a hypothesis whose premises include empirical data. To be precise, one can deductively falsify a hypothesis based on experimental results, but there can be no instance in which one can deductively verify it.

1.6.4.2 Reductio Ad Absurdum

Another type of logical argument, important to null hypothesis testing, is *reductio ad absurdum*, which means "reduction to the absurd." Here, one tries to support a hypothesis, H, by formulating a hypothesis that represents not-H (i.e., the opposite of H), and then disproving not-H.

EXAMPLE 1.2

One of the most famous examples of *reductio ad absurdum* is a proof, credited to Pythagoras (570–495 BC), which demonstrates that $\sqrt{2}$; is irrational. I will break his arguments into eight steps.

1. Assume there is a rational number whose square is equal to two.
2. This number must be a fraction since there is no integer whose square equals two. Let this fraction be reduced to its lowest common factor. We have

$$(a/b)^2 = 2 \quad \text{and, as a result} \quad a^2 = 2b^2$$

3. a^2 must be even because the product of two and any other number must be even. In addition, a must be even because the square of any even number must be even, while the square of any odd number must be odd.
4. Let $a = 2c$ to remind us that a is even. We have

$$a = 2c \quad \text{and} \quad a^2 = 4c^2$$

* Another logically correct form of deduction is called *affirming the antecedent*. Here is its general form:

If H, then I } Premise 1
H } Premise 2
I } Conclusion

Like all logically correct deductive arguments, any statement can be selected for H and I, and the conclusions will be true if the premises are true. In contrast, with affirming the consequence we have

If H, then I } Premise 1
I } Premise 2
H } Conclusion

As with argument 1.2, the statement above is deductively invalid.This can be demonstrated by substituting statements for H and I and constructing an argument with true premises and a false conclusion.

If Portland is in Washington state, then it is in the Pacific Northwest } Premise 1
Portland is in the Pacific Northwest } Premise 2
Portland is in Washington state } Conclusion

5. Substituting back into our earlier equation, we have

$$4c^2 = 2b^2 \quad \text{and} \quad 2c^2 = b^2$$

6. From step 3, we see that b must also be even.
7. But if a and b are both divisible by two, then they have not been reduced to their smallest common factor, and this contradicts the assumption in step two.
8. We conclude that it is not possible to have a rational number whose square equals two. That is, we conclude $\sqrt{2}$ is irrational.

This argument follows the form of *reductio ad absurdum*. To prove that $\sqrt{2}$ is irrational, we assume that the elements making up $\sqrt{2}$ are rational. Because we contradict this assumption, we then disprove the hypothesis that $\sqrt{2}$ is rational.*

The *asbio* package provides several interactive logic worksheets illustrating important types of logical and fallacious arguments. These can be accessed from the `book.menu()` or by typing `see.logic()` in the **R** console after installing and loading **asbio**.

1.7 Variability and Uncertainty in Investigations

The impossibility of proving a deductive argument based on empirical evidence is due to the variability of natural systems, and the fact that all possible data concerning these systems cannot be collected (Section 1.6). If natural phenomena in the universe never varied, then conclusive empirical proof would be possible. In fact, one observation would be sufficient to establish truth/falsehood. However, this does not happen.

In the elk experiment from Example 1.1, there are several potential sources of confounding variability. One possibility is that willow survivorship does not respond linearly to the presence of elk. For example, it may take a large number of elk to cause any sort of decrease in willow seedling recruitment. If this is true, then a researcher must be conscious of this particular relationship to be able to detect the effect of elk. A second confounding outcome will almost certainly occur. This is that replicates in the exclosure and control groves will vary. For instance, it is possible that elk will graze only at some control groves and ignore other control groves to avoid a pack of wolves that move into the area. Another possibility is that some exclosure groves will be unknowingly located on soils that are particularly nutrient poor. As a result, these groves will not respond with increased recruitment even though elk grazing is prevented. In both cases, the negative effect of the elk on the willows will be obscured by sample variability.

All empirical data describing the universe will contain variability that can confound decisions.† This extends to even the most venerable scientific "truisms." For instance, it is

* It is interesting that the followers of Pythagoras, who believed that numbers were sacred, withheld the blasphemous information that $\sqrt{2}$ is irrational.

† The natural propensity of experimental outcomes to vary from sample to sample has been given the rather unfortunate name, *sampling error*. The capacity to measure g (or anything else) will also be limited because of *measurement error*. For instance, currently the most precise tool for measuring g is the absolute atom gravimeter (Müller et al. 2008). While these devices are incredibly accurate (they give readings close to the "true" value of g) and precise (i.e., they give very similar repeated measures), they still exhibit a degree of uncertainty $\approx 1.3 \times 10^{-9}\ g$. Sampling error, measurement error, accuracy, and precision are addressed repeatedly in this text.

often taken for granted that the acceleration of objects due to gravity, g, near the earth's surface = 9.8 $\mathrm{m \cdot s^{-2}}$. However, we now know that g varies with the local density of the earth's surface and the speed of the earth's rotation, which in turn varies with latitude (de Angelis et al. 2009).

It has been suggested that the laws of physics (e.g., laws of electrostatics and thermodynamics) differ from those of biology since they provide tidy deterministic models and often address objects (such as electrons) that are nonvariable (Pastor 2008). However, the empirical behavior of matter at subatomic scales (e.g., electrons) is *not* deterministic. In fact the *Heisenberg uncertainty principle* stipulates that it is impossible, even in theory, to simultaneously know the exact location and momentum of a subatomic particle (Heisenberg 1927). Indeed, the more precisely one attribute is measured, the less precisely the other can be measured. As a result, field of quantum mechanics defines *stochastic* (nondeterministic) models to describe the workings of subatomic particles.* Stochastic processes also underlie much larger presumably deterministic systems, including the orbits of moons and planets in the solar system (Peterson 1993).†

Scientific theories that have withstood repeated testing (e.g., the theory of evolution and the theory of relativity) can be trusted as objective, realistic, and corroborated pictures of the universe. However, they should not be treated dogmatically (a behavior called *scientism*). Scientists and others should recognize that the variability implicit in natural phenomena (and measurements of phenomena) means that absolute proof for a hypothesis based on empirical data can never be obtained. Acknowledgment of this state of affairs allows continual refutation and revision of hypotheses, and the continual development and improvement of scientific ideas. Empirical science, of course, is not indifferent to truth. It is more correct, however, to view it as *truth-directed*. That is, it seeks to know, but acknowledges at the outset that it can never know for certain.‡

In conclusion, it is important to emphasize two things. First, the refutation of a hypothesis may provide valuable information about the study system by allowing a researcher to eliminate incorrect lines of reasoning, and clarify its true nature. Second, even though empirical evidence cannot be used to prove that a hypothesis is true, it can be used to corroborate a hypothesis. Thus, while nonrefutation is inconclusive, it often provides valuable supporting evidence. The continual accumulation of such evidence, and a lack of negating evidence, may indicate a very strong statement about the natural world, allowing the development of scientific theories and laws (Ayala et al. 2008).

The importance of trustworthy tools to measure the empirical support for hypotheses is obvious. Along with dependable field and laboratory instruments, these include *mathematical models* (descriptions based on mathematical concepts), particularly those associated with statistics.

---

* Albert Einstein (1879–1955) found the lack of determinism in quantum mechanics to be unsettling, giving rise to his famous quote: "God does not play dice with the universe."

† The orbits of planets are perhaps better described as *chaotic* not *stochastic* (Peterson 1993, Denny and Gaines 2002). While stochastic simply means that a process is nondeterministic, chaotic refers to a process that is deterministic, but very sensitive to initial conditions that determine its outcome. As a practical matter, chaotic processes are often effectively modeled with stochastic tools like probability.

‡ A view of science as truth-directed is in accordance with *logical positivism* (Okasha 2002) and subsequently, with *scientific realism*. This outlook has been opposed, notably by Kuhn (1963) who insisted that scientific truth is relative to the current scientific paradigm (Laudan 1977). Famous logical positivists included Moritz Schlick (1882–1936), Ernst Mach (1838–1916), and the young Ludwig Wittgenstein (1889–1951). Important scientific realists include Karl Popper (1902–1994) and Richard Boyd (1942–).

1.8 Science and Statistics

It is apparent that empirical scientists are faced with a dilemma. We need to objectively quantify evidence supporting our hypotheses, but we must acknowledge at the outset that the variability of the world makes it impossible to ever prove a hypothesis. A reasonable response to this problem is statistics. Statistical models explicitly acknowledge uncertainty, and through mathematics, allow us to estimate variability and evaluate outcomes probabilistically.

The word *statistics* is conventionally defined as "a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data."* To empirical scientists, however, this definition seems insufficient. For these individuals statistics allows (1) the expression and testing of research hypotheses within the framework of the scientific method, (2) the use of mathematics to make precise and objective inferences, and (3) the quantification of uncertainty implicit to natural systems. As Ramsey and Schafer (1997) note: "statistics is like grout—the word feels decidedly unpleasant in the mouth, but it describes something essential for holding a mosaic in place."

1.9 Statistics and Biology

The development of modern statistical methods has been largely due to the efforts of biologists or of mathematicians who have dealt almost exclusively with biological systems. Francis Galton (1822–1911) was the first cousin of Darwin, and the inventor of regression analysis (Chapter 9). Karl Pearson (1857–1936) invented the correlation coefficient (Chapter 8), the χ^2 contingency test (Chapter 11), and was one of the first practical users of probability distributions (Chapter 3). Together, Galton and Pearson founded the journal *Biometrika* to define distributional parameters for biological populations with the ambitious goal of quantifying the process of evolution as the result of natural selection (Salsburg 2001). William Sealy Gosset (1876–1937) developed a number of important ideas in mathematical statistics, including the *t*-distribution (Chapter 6), while working as a biometrist for the Guinness brewing company. Roland A. Fisher (1890–1962) has been hailed by mathematical historians as single-handedly creating the "foundations for modern statistical science" (Hald 1998). His achievements include the invention of maximum likelihood estimation (Chapter 4), the *P*-value (Chapter 6), the null hypothesis (Chapter 6), and the analysis of variance (Chapter 10). Fisher principally analyzed and interpreted biological data during his long career, and also made important scientific contributions as a geneticist and evolutionary biologist. Indeed, he has been described as "the greatest of Darwin's successors" (Dawkins 1995). Jerome Cornfield (1912–1979) was not only a developer of ideas associated with Bayesian inference (Chapter 6) and experimental design (Chapter 7) but also made important contributions to epidemiology, including the identification of carcinogens. A.W.F. Edwards (1935–), whose books have clarified the importance of likelihood in inferential procedures, has worked principally as a geneticist and evolutionary biologist. The list goes on and on (Table 1.2).

* *Merriam Webster's New Collegiate Dictionary*, 11th edition.

atistics for Biologists Using R

ıma. We need to objectively
acknowledge at the outset
:ove a hypothesis. A reason-
:plicitly acknowledge uncer-
ility and evaluate outcomes

f mathematics dealing with
ısses of numerical data."* To
ıt. For these individuals sta-
ıeses within the framework
precise and objective infer-
ral systems. As Ramsey and
lecidedly unpleasant in the
aic in place."

ely due to the efforts of biol-
ely with biological systems.
d the inventor of regression
relation coefficient (Chapter
t practical users of probabil-
ınded the journal *Biometrika*
ıs with the ambitious goal
al selection (Salsburg 2001).
ortant ideas in mathematical
king as a biometrist for the
been hailed by mathemati-
modern statistical science"
mum likelihood estimation
ıpter 6), and the analysis of
preted biological data dur-
ibutions as a geneticist and
e greatest of Darwin's suc-
ot only a developer of ideas
ntal design (Chapter 7) but
the identification of carcin-
importance of likelihood in
and evolutionary biologist.

TABLE 1.2
An Incomplete Hagiography of Important 20th- and 21st-Century Statisticians with Strong Biological Connections

Name	Statistical Contributions	Biological Connections/Contributions
Karl Pearson (1857–1936)	With *Francis Galton* (1822–1911) invented the correlation coefficient. Played a key role in the early development of regression analysis. Invented the chi-squared test, and the method of moments for deriving estimators.	Founded the journals *Biometrika* and *Annals of Human Genetics*. Involved in epidemiology and medicine. Evaluated evidence for the theory of evolution with species probability distributions.
R.A. Fisher (1890–1962)	Enormously important figure in the development of mathematical and applied statistics. Contributions include the *F*-distribution, ANOVA, maximum likelihood estimation, and many other approaches.	A leading figure in evolutionary ecology, quantitative population genetics, and population ecology. Biostatistician at the Rothamsted agricultural experimental station.
Jerzy Neyman (1894–1981)	With K. Pearson's son *Egon Pearson* (1895–1980), and battling R.A. Fisher, developed a framework for null hypothesis testing, and the concepts of type I and type II error.	Founded biometric laboratories in Poland. Established methods for testing medicines still used today by the United States Food and Drug Administration.
W.S. Gosset (1876–1937)	Proposed and developed the *t*-distribution. To assuage fears of his employer (Guinness) that he would give away brewery secrets, Gosset published under the pseudonym "student."	For many years studied methods to maximize barley yield.
A.N. Kolmogorov (1903–1987)	Made large contributions to the understanding of stochastic systems, particularly Markov processes. Created applied statistical innovations, including the Kolmogorov–Smirnov test.	Generalized and expanded the Lotka–Volterra model of predator–prey interactions, and postulated models for the structure of the brain.
G.W. Snedecor (1881–1974)	With *W.G. Cochran* (1909–1980) wrote the most frequently cited statistical text in existence: "Statistical Methods." The *F*-distribution is often called Snedecor's *F*-distribution to acknowledge the fact that he handcrafted the first *F* cumulative distribution tables.	Deeply involved in agricultural research. Published influential biometric texts including: "Statistical Methods Applied to Experiments in Agriculture and Biology."
Gertrude Cox (1900–1978)	With *David Box* (1924–) developed maximum likelihood methods for optimal variable transformations in linear models. With W.G. Cochran published an important book on experimental design.	Editor of the *Biometrics Bulletin* and of *Biometrics*, and founder of the International Biometric Society.
J. Cornfield (1912–1979)	Contributed to applications in Bayesian inference. Formalized the importance of observational studies. Played a role in the early development of linear programming.	Strong research interests in human health, particularly the effects of smoking. Developed important epidemiologic experimental approaches.
John Nelder (1924–2010)	With *Peter McCullagh*, developed the unifying concept of generalized linear models (GLMs). Proponent for the inferential use of likelihood.	Strong interests in ornithology and evolutionary biology. Wrote a book on computational biology. Served as head of biostatistics at the Rothamsted experimental station.
A.W.F. Edwards (1935–)	Strong proponent of the inferential use of likelihood. Lucid statistical writer.	Geneticist and evolutionary biologist. Developed methods for quantitative phylogenetic analysis.

Note: Statistical terms in the table are explained in Chapters 2 through 11.

The need for biologists to be able to understand and apply statistical and mathematical models has never been greater. This is true for two reasons. First, biologists are faced with the tremendous responsibility of being able to accurately, precisely, and objectively describe the world. This includes the complex global problems we currently face, for example, global climate change, epidemics, human aging and disease, alteration of nutrient cycles, increased resource demands from overpopulation and cultural development, habitat fragmentation, loss of biodiversity, acidification of rain, and ozone thinning. Second, developments in statistical theory and computer technology have provided a bewildering new array of analytical tools for biologists. If handled correctly and intelligently, these tools provide an unprecedented medium to describe our world, and address our problems. Conversely, if these tools are handled inappropriately, they will produce ambiguous or incorrect results, and give rise to distrust of scientific claims.

1.10 Summary

This chapter provides a motivation for the development of the remainder of the book, and a philosophical and historical context for both science and statistics. Scientists seek to provide truth-directed descriptions of the world. To accomplish this, scientific descriptions are based on three principles: objectivity, realism, and communalism. The scientific method is built on hypotheses, that is, statements describing the universe, which serve as starting points for investigations. Scientific hypotheses have two characteristics: testability and (arguably) falsifiability. Testability is explicit to the scientific method, while falsifiability requires the possibility of refutation. The approaches used by scientists to understand the natural world are based on precepts of logic. Two modes of logic are possible: induction and deduction. The conclusions from an inductive argument contain information not present, even implicitly, in the premises. Conversely, a deductive conclusion is contained entirely in the premises. All natural systems vary, including those studied by biologists. Thus, information about those systems will be incomplete. Because of this, absolute proof of the truth of a hypothesis concerning the natural world is impossible. In the face of these issues, biologists require a way to describe the trends and variability in the systems they study, and to quantify evidence concerning their hypotheses (despite this variability). A rational solution to this problem is the discipline of statistics.

EXERCISES

1. Distinguish "empirical science" and "nonempirical science."
2. Define the terms "objectivity," "scientific realism," and "communalism."
3. Describe two characteristics of a scientific hypothesis. Given these characteristics, list three hypotheses that could fall under the heading "scientific hypothesis" and three that could fall under the heading "nonscientific hypothesis." Justify your reasoning.
4. Describe similarities and differences in the views on falsifiability of Popper, Kuhn, and Lakatos.
5. How does the first premise of scientific realism given in this chapter provide a connection to falsificationism?

statistical and mathemati-
s. First, biologists are faced
ý, precisely, and objectively
we currently face, for exam-
ease, alteration of nutrient
cultural development, habi-
ıd ozone thinning. Second,
ave provided a bewildering
ctly and intelligently, these
, and address our problems.
vill produce ambiguous or

---

he remainder of the book,
ıd statistics. Scientists seek
lish this, scientific descrip-
mmunalism. The scientific
he universe, which serve as
o characteristics: testability
ic method, while falsifiabil-
by scientists to understand
of logic are possible: induc-
ent contain information not
ive conclusion is contained
those studied by biologists.
cause of this, absolute proof
impossible. In the face of
d variability in the systems
ses (despite this variability).
s.

ce."
ommunalism."
ven these characteristics,
cientific hypothesis" and
ıypothesis." Justify your

fiability of Popper, Kuhn,

n this chapter provide a

6. Define the terms "fallacious" and "logically correct."
7. Distinguish inductive and deductive reasoning. Give an example of each in the context of a biological investigation.
8. Create an example of a logically correct argument with false premises, and a fallacious argument with true premises.
9. Give examples of *modus tollens* (*denying the consequent*) and *affirming the consequent* arguments. Explain why the latter is logically incorrect.
10. Give an example of *reduction ad absurdum*.
11. Go through the interactive logic worksheet provided by installing and loading the *asbio* package and typing `see.logic()`. Come up with your own syllogistic arguments illustrating *modus ponens, post hoc ergo propter hoc*, and confusing cause and effect. Indicate whether these arguments are logical or fallacious.
12. According to the author, why should science be described as "truth-directed"?
13. Why does the field of statistics fit well with the goals and realities of empirical science?
14. Choose an individual from Table 1.2 and create an expanded biography, or prepare a brief summary for another biostatistician.