1 Introduction

1.1 UNDERSTANDING THE WORLD THROUGH EVIDENCE-BASED REASONING

We can try to understand the world in various ways, an obvious one being the employment of empirical methods for gathering and analyzing various forms of evidence about phenomena, events, and situations of interest to us. This will include work in all of the sciences, medicine, law, intelligence analysis, history, political affairs, current events, and a variety of other contexts too numerous to mention. In the sciences, this empirical work will involve both experimental and nonexperimental methods. In some of these contexts, notably in the sciences, we are able to devise mathematical and logical models that allow us to make inferences and predictions about complex matters of interest to us. But in every case, our understanding rests on our knowledge of the properties, uses, discovery, and marshaling of evidence. This is why we begin this book with a careful consideration of reasoning based on evidence.

1.1.1 What Is Evidence?

You might think this question is unnecessary since everyone knows what evidence is. However, matters are not quite that simple, since the term *evidence* is not so easy to define and its use often arouses controversy. One problem with the definition of evidence is that several other terms are often used synonymously with it, when in fact there are distinctions to be made among these terms that are not always apparent. Quite unnecessary controversy occurs since some believe that the term *evidence* arises and has meaning only in the field of law.

Consulting a dictionary actually does not assist us much in defining the term. For example, look at the *Oxford English Dictionary* under the term *evidence* and you will be led in a circle; *evidence* is ultimately defined as being evidence.

A variety of terms are often used as synonyms for the term *evidence*: *data*, *items of information*, *facts*, and *knowledge*. When examined carefully, there are some valid and important distinctions to be made among these terms, as we will now discuss.

1.1.2 Evidence, Data, and Information

Consider the terms *data* and *items of information*.

Data are uninterpreted signals, raw observations, measurements, such as the number 6, the color “red,” or the sequence of dots and lines “. . .–. . .”.

Information is data equipped with meaning provided by a certain context, such as "6 AM," "red traffic light," "red tomato," or the "S O S" emergency alert.

Untold trillions of data and items of information exist that will almost certainly never become evidence in most inferences. Here's an item of information for you: Professor Schum has a long and steep driveway in front of his house that makes shoveling snow off of it very difficult in the winter. Can you think of any situation in which this item of information could become evidence? About the only matter in which this information could become interesting evidence involves the question: Why did Schum and his wife, Anne, ever purchase this house in the first place? As we will discuss, *items of information become evidence only when their relevance is established regarding some matter to be proved or disproved.*

1.1.3 Evidence and Fact

Now consider the term *fact*; there are some real troubles here as far as its relation to the term *evidence* is concerned. How many times have you heard someone say, "I want all the facts before I draw a conclusion or make a decision," or, "I want to know the facts in this matter"? The first question is easily answered: We will never have all the facts in any matter of inferential interest. Answers to the second question require a bit of careful thought. Here is an example of what is involved:

Suppose we are police investigators interviewing Bob, who is a witness of a car accident that just happened at a particular intersection. Bob tells us that the Ford car did not stop at the red light signal. Now we regard it as fact that Bob gave us this information: We all just heard him give it to us. But whether the Ford car did not stop at the red light *is only an inference and is not a fact*. This is precisely why we need to distinguish carefully between an *event* and *evidence* for this event.

Here is what we have: Bob has given us evidence E*, saying that event E occurred, where E is the event that the Ford car did not stop at the red light signal. Whether this event E did occur or not is open to question and depends on Bob's competence and credibility. If we take it as *fact* that event E did occur, just because Bob said it did, we would be overlooking the *believability* foundation for any inference we might make from his evidence E*. Unfortunately, it so often happens that people regard the events reported in evidence as being facts when they are not. Doing this suppresses all uncertainties we may have about the source's credibility and competence if the evidence is testimonial in nature. We have exactly the same concerns about the credibility of tangible evidence. For example, we have been given a tangible object or an image as evidence E* that we believe reveals the occurrence of event E. But we must consider whether this object or image is authentic and it is what we believe it to be. In any case, the events recorded in evidence can be regarded as facts only if provided by perfectly credible sources, something we almost never have. As another example, any information we find on the Internet should be considered as only a claim by its source rather than as fact, that is, as *evidence* about a potential fact rather than a *fact*.

1.1.4 Evidence and Knowledge

Now consider the term *knowledge* and its relation with evidence. Here is where things get interesting and difficult. As you know, *the field of epistemology is the study of knowledge,* *what we believe it may be, and how we obtain it.* Two questions we would normally ask regarding what Bob just told us are as follows:

- Does Bob really *know* what he just told us, that the Ford car did not stop at the red light signal?
- Do we ourselves then also *know*, based on Bob's testimony, that the Ford car did not stop at the red light signal?

Let's consider the first question. For more than two millennia, some very learned people have troubled over the question: What do we mean when we say that person A *knows* that event B occurred? To apply this question to our source Bob, let's make an assumption that will simplify our answering this question. Let's assume that Bob is a *competent* observer in this matter. Suppose we have evidence that Bob was actually himself at the intersection when the accident happened. This is a major element of Bob's competence. Bob's credibility depends on different matters, as we will see.

Here is what a standard or conventional account says about whether Bob knows that the car did not stop at the red light signal. First, here is a general statement of the standard account of knowledge: *Knowledge is justified true belief*. Person $\mathcal{A}$ knows that event B occurred if:

- Event B did occur [True];
- $\mathcal{A}$ got nondefective evidence that B occurred [Justified]; and
- $\mathcal{A}$ believed this evidence [Belief].

This standard analysis first says that event B must have occurred for $\mathcal{A}$ to have knowledge of its occurrence. This is what makes $\mathcal{A}$'s belief true. If B did not occur, then $\mathcal{A}$ could not know that it occurred. Second, $\mathcal{A}$'s getting nondefective evidence that B occurred is actually where $\mathcal{A}$'s competence arises. $\mathcal{A}$ could not have gotten any evidence, defective or nondefective, if $\mathcal{A}$ was not where B could have occurred. Then, $\mathcal{A}$ believed the evidence $\mathcal{A}$ received about the occurrence of event B, and $\mathcal{A}$ was justified in having this belief by obtaining nondefective evidence of B's occurrence.

So, in the case involving Bob's evidence, Bob knows that the Ford car did not stop at the red light signal if:

- The car did not stop at the red light signal,
- Bob got nondefective evidence that the car did not stop at the red signal, and
- Bob believed this evidence.

If all of these three things are true, we can state on this standard analysis that Bob knows that the Ford car did not stop at the red light signal.

Before we proceed, we must acknowledge that this standard analysis has been very controversial in fairly recent years and some philosophers claim to have found alleged paradoxes and counterexamples associated with it. Other philosophers dispute these claims. Most of the controversy here concerns the justification condition: What does it mean to say that A is justified in believing that B occurred? In any case, we have found this standard analysis very useful as a heuristic in our analyses of the credibility of testimonial evidence.

But now we have several matters to consider in answering the second question: Do we ourselves also *know*, based on Bob's testimony, that the Ford car did not stop at the red light signal? The first and most obvious fact is that we do not know the extent to which any of the three events just described in the standard analysis are true. We cannot get inside

Bob's head to obtain necessary answers about these events. Starting at the bottom, we do not know for sure that Bob believes what he just told us about the Ford car not stopping at the red light signal. This is a matter of Bob's *veracity* or *truthfulness*. We would not say that Bob is being truthful if he told us something he did not believe.

Second, we do not know what sensory evidence Bob obtained on which to base his belief and whether he based his belief at all on this evidence. Bob might have believed that the Ford car did not stop at the red light signal either because he expected or desired this to be true. This involves Bob's *objectivity* as an observer. We would not say that Bob was objective in this observation if he did not base his belief on the sensory evidence he obtained in his observation.

Finally, even if we believe that Bob was an objective observer who based his belief about the accident on sensory evidence, we do not know how good this evidence was. Here we are obliged to consider Bob's sensory *sensitivities or accuracy in the conditions under which Bob made his observations*. Here we consider such obvious things as Bob's visual acuity. But there are many other considerations, such as, "Did Bob only get a fleeting look at the accident when it happened?" "Is Bob color-blind?" "Did he make this observation during a storm?" and, "What time of day did he make this observation?" For a variety of such reasons, Bob might simply have been mistaken in his observation: The light signal was not red when the Ford car entered the intersection.

So, what it comes down to is that the extent of our knowledge about whether the Ford car did not stop at the red light signal, based on Bob's evidence, depends on these three attributes of Bob's credibility: his veracity, objectivity, and observational sensitivity. We will have much more to say about assessing the credibility of sources of evidence, and how Disciple-EBR can assist you in this difficult process, in Section 4.7 of this book.

Now, we return to our role as police investigators. Based on evidence we have about Bob's competence and credibility, suppose we believe that the event he reports did occur; we believe that "E: The Ford car did not stop at the red light signal," did occur. Now we face the question: So what? Why is knowledge of event E of importance to us? Stated more precisely: How is event E relevant in further inferences we must make? In our investigation so far, we have other evidence besides Bob's testimony. In particular, we observe a Toyota car that has smashed into a light pole at this intersection, injuring the driver of the Toyota, who was immediately taken to a hospital. In our minds, we form the tentative chain of reasoning from Figure 1.1.

This sequence of events, E ➔ F ➔ G ➔ H, is a relevance argument or chain of reasoning whose links represent sources of doubt interposed between the evidence E* and the hypothesis H. An important thing to note is that some or all of these events may not be true. Reducing our doubts or uncertainties regarding any of these events requires a variety

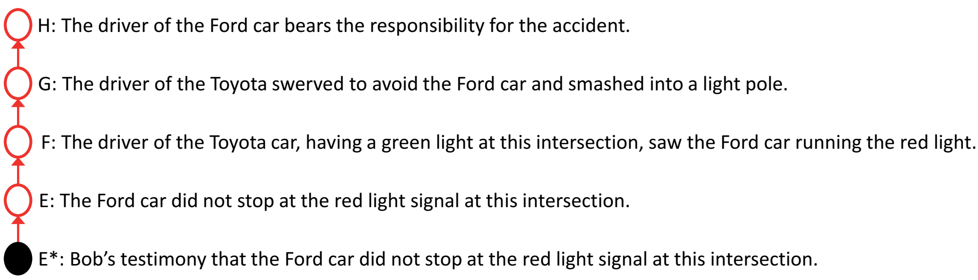

Figure 1.1. Tentative chain of reasoning.

of additional evidence. The extent of our knowledge about the relative probabilities of our final hypothesis depends on the believability of our evidence and on the defensibility and strength of our relevance arguments, as discussed in Section 2.2. The whole point here is that the relation between evidence and knowledge is not a simple one at all.

1.1.5 Ubiquity of Evidence

Finally, we must consider the controversy over the use of the term *evidence* instead of the other terms we just examined. The mistake made by some people is to consider that evidence concerns only objects, testimony, or other items introduced in a court trial. This controversy and confusion has been recognized by eminent evidence scholars in the field of law. For example, in his marvelous book *Evidence, Proof, and Facts: A Book of Sources*, Professor Peter Murphy (2003, p. 1) notes the curious fact that the term evidence is so commonly associated only with the field of law:

> The word "evidence" is associated more often with lawyers and judicial trials than with any other cross-section of society or form of activity. . . . In its simplest sense, evidence may be defined as any factual datum which in some manner assists in drawing conclusions, either favorable or unfavorable, to some hypothesis whose proof or refutation is being attempted.

Murphy notes that this term is appropriate in any field in which conclusions are reached from any relevant datum. Thus, physicians, scientists of any ilk, historians, and persons of any other conceivable discipline, as well as ordinary persons, use evidence every day in order to draw conclusions about matters of interest to them.

We believe there is a very good reason why many persons are so often tempted to associate the term *evidence* only with the field of law. It happens that the Anglo-American system of laws has provided us with by far the richest legacy of experience and scholarship on evidence of any field known to us. This legacy has arisen as a result of the development of the adversarial system for settling disputes and the gradual emergence of the jury system, in which members of the jury deliberate on evidence provided by external witnesses. This legacy has now been accumulating over at least the past six hundred years (Anderson et al., 2005).

Evidence-based reasoning involves abductive, deductive, and inductive (probabilistic) reasoning. The following sections briefly introduce them.

1.2 ABDUCTIVE REASONING

1.2.1 From Aristotle to Peirce

Throughout history, some of the greatest minds have tried to understand the world through a process of discovery and testing of hypotheses based on evidence. We have found the metaphor of an *arch of knowledge* to be very useful in summarizing the many ideas expressed over the centuries concerning the generation of new thoughts and new evidence. This metaphor comes from the work of the philosopher David Oldroyd in a valuable work having this metaphor as its title (Oldroyd, 1986). Figure 1.2 shows this metaphor applied in the context of science. Based upon existing records, it seems that

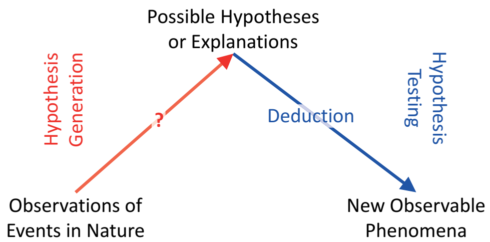

Figure 1.2. The "arch of knowledge" in science.

Aristotle (384 BC–322 BC) was the first to puzzle about the generation or discovery of new ideas in science. From sensory observations, we generate possible explanations, in the form of hypotheses, for these observations. It was never clear from Aristotle's work what label should be placed on the upward, or discovery-related, arm of the arch in Figure 1.2. By some accounts, Aristotle's description of this act of generating hypotheses is called "intuitive induction" (Cohen and Nagel, 1934; Kneale, 1949). The question mark on the upward arm of the arch in Figure 1.2 simply indicates that there is still argument about what this discovery-related arm should be called. By most accounts, the downward arm of the arch concerns the deduction of new observable phenomena, assuming the truth of a generated hypothesis (Schum, 2001b).

Over the millennia since Aristotle, many people have tried to give an account of the process of discovering hypotheses and how this process differs from ones in which existing hypotheses are justified. Galileo Galilei (1564–1642) thought that we "reason backward" inductively to imagine causes (hypotheses) from observed events, and we reason deductively to test the hypotheses. A similar view was held by Isaac Newton (1642–1727), John Locke (1632–1704), and William Whewell (1794–1866). Charles S. Peirce (1839–1914) was the first to suggest that new ideas or hypotheses are generated through a different form of reasoning, which he called *abduction* and associated with imaginative reasoning (Peirce, 1898; 1901). His views are very similar to those of Sherlock Holmes, the famous fictional character of Conan Doyle (Schum, 1999).

1.2.2 Peirce and Sherlock Holmes on Abductive Reasoning

Until the time of Peirce, most persons interested in discovery and investigation supposed that the discovery-related arms of the arch in Figure 1.2 involved some form of inductive reasoning that proceeds from particulars, in the form of observations, to generalities, in the form of hypotheses. But inductive reasoning is commonly associated with the process of justifying or trying to prove existing hypotheses based on evidence. The question remains: Where did these hypotheses come from? Pondering such matters, Peirce relied on a figure of reasoning he found in the works of Aristotle. The reasoning proceeds as follows:

- If H were true, then E, F, and G would follow as a matter of course.
- But E, F, and G have been observed.
- Therefore, we have reason to believe that H *might possibly* be true.

As an illustration, let us assume that we observe E*: "Smoke in the East building" (E* being evidence that event E occurred).

Based on our prior knowledge of contexts in which things like E: "Smoke in a building" have occurred, we say: "Whenever something like H: 'There is fire in a building' has occurred, then something like E: 'Smoke in the building' has also occurred." Thus, there is reason to suspect that H: "There is fire in the East building" may explain the occurrence of the clue E*: "Smoke in the East building." In other words, the clue E* points to H as a possible explanation for its occurrence.

To summarize:

- We observe smoke in the East building.
- Fire causes smoke.
- We hypothesize that there is a fire in the East building.

Peirce was unsure about what to call this form of reasoning. At various points in his work, he called it "abduction," "retroduction," and even just "hypothesis" (Pierce, 1898; 1901).

The essential interpretation Peirce placed on the concept of abduction is illustrated in Figure 1.3. He often used as a basis for his discussions of abduction the observation of an anomaly in science. Let us suppose that we already have a collection of prior evidence in some investigation and an existing collection of hypotheses H_1, H_2, . . . , H_n. To varying degrees, these *n* hypotheses explain the evidence we have so far. But now we make an observation E* that is embarrassing in the following way: We take E* seriously, but we cannot explain it by any of the hypotheses we have generated so far. In other words, E* is an anomaly. Vexed by this anomaly, we try to find an explanation for it. In some cases, often much later when we are occupied by other things, we experience a "flash of insight" in which it occurs to us that a new hypothesis H_{n+1} could explain this anomaly E*. It is these "flashes of insight" that Peirce associated with abduction. Asked at this moment to say exactly how H_{n+1} explains E*, we may be unable to do so. However, further thought may produce a chain of reasoning that plausibly connects H_{n+1} and E*. The reasoning might go as follows:

- I have evidence E* that event E happened.
- If E did happen, then F might be true.
- If F happened, then G might be true.
- And if G happened, then H_{n+1} might be true.

It is possible, of course, that the chain of reasoning might have started at the top with H_{n+1} and ended at E*. This is why we have shown no direction on the links between E* and H_{n+1} in Figure 1.3.

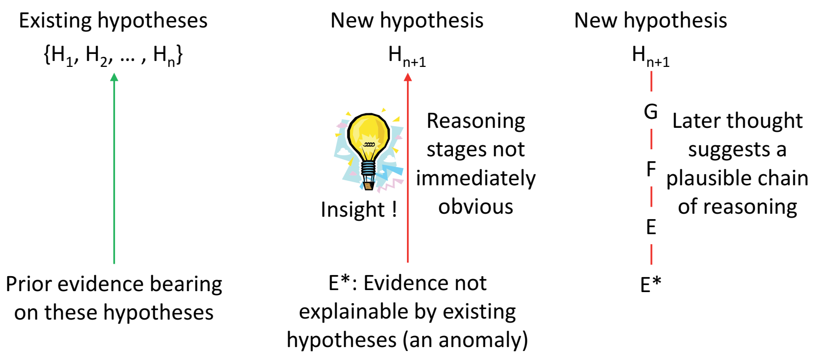

Figure 1.3. Peirce's interpretation of abductive reasoning.

But our discovery-related activities are hardly over just because we have explained this anomaly. Our new hypothesis H_{n+1} would not be very appealing if it explained only anomaly E*. Figure 1.4 shows the next steps in our use of this new hypothesis. We first inquire about the extent to which it explains the prior evidence we collected before we observed E*. An important test of the suitability of the new hypothesis H_{n+1} involves asking how well this new hypothesis explains other observations we have taken seriously. This new hypothesis would be especially valuable if it explains our prior evidence better than any of our previously generated hypotheses. But there is one other most important test of the adequacy of a new hypothesis H_{n+1}: How well does this new hypothesis suggest new potentially observable evidence that our previous hypotheses did not suggest? If H_{n+1} would be true, then B, I, and K would also be true; and if B would be true, then C would be true. Now if C would be true, then we would need to observe D.

In the illustrations Peirce used, which are shown in Figures 1.3 and 1.4, we entered the process of discovery at an intermediate point when we already had existing hypotheses and evidence. In other contexts, we must of course consider abductive reasoning from the beginning of an episode of fact investigation when we have no hypotheses and no evidence bearing on them. Based on our initial observations, by this process of abductive or insightful reasoning, we may generate initial guesses or hypotheses to explain even the very first observations we make. Such hypotheses may of course be vague, imprecise, or undifferentiated. Further observations and evidence we collect may allow us to make an initial hypothesis more precise and may of course suggest entirely new hypotheses.

It happens that at the very same time Peirce was writing about abductive reasoning, insight, and discovery, across the Atlantic, Arthur Conan Doyle was exercising his fictional character Sherlock Holmes in many mystery stories. At several points in the Sherlock Holmes stories, Holmes describes to his colleague, Dr. Watson, his inferential strategies during investigation. These strategies seem almost identical to the concept of abductive reasoning described by Peirce. Holmes did not, of course, describe his investigative reasoning as abductive. Instead, he said his reasoning was “backward," moving from his observations to possible explanations for them. A very informative and enjoyable collection of papers on the connection between Peirce and Sherlock Holmes appears in the work of Umberto Eco and Thomas Sebeok (1983). In spite of the similarity of Peirce's and Holmes's (Conan Doyle's) views of discovery-related reasoning, there is no evidence that Peirce and Conan Doyle ever shared ideas on the subject.

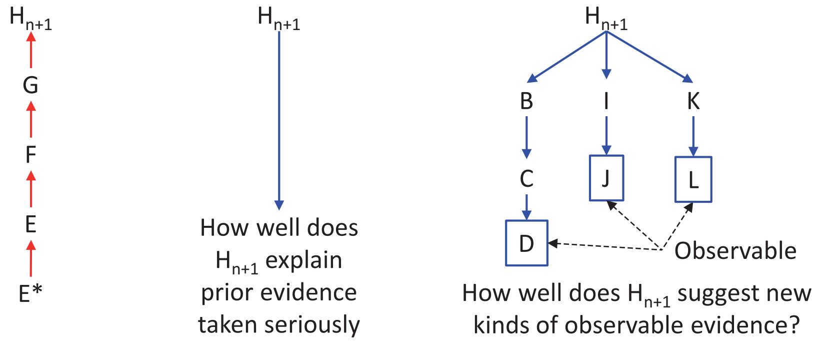

Figure 1.4. Putting an abduced hypothesis to work.

1.3 PROBABILISTIC REASONING

A major trouble we all face in thinking about probability and uncertainty concerns the fact that the necessity for probability calculations, estimations, or judgments arises in different situations. In addition, there are many different attributes of our judgments that we would like to capture in assessments of uncertainty we are obliged to make. There are situations in which you can estimate probabilities of interest by counting things. But there are many other situations in which we have uncertainty but will have nothing to count. These situations involve events that are singular, unique, or one of a kind. In the following, we will briefly discuss several alternative views of probability, starting with two views of probability that involve processes in which we can obtain probabilities or estimates of them by enumerative or counting processes.

1.3.1 Enumerative Probabilities: Obtained by Counting

1.3.1.1 Aleatory Probability

According to Laplace (1814, p. cv), "probability theory is nothing but common sense reduced to calculation." There are two conceptions of probability that involve counting operations. The first is termed *aleatory probability*. This term has its roots in the Latin term *alea,* meaning chance, game of chance, or devices such as dice involved in games of chance. Games of chance have two important ground rules:

- There is a finite number n(S) of possible outcomes
- All outcomes in S are assumed to have equal probability

For example, in a game involving a pair of fair six-sided dice, where we roll and add the two numbers showing up, there are thirty-six ways in which the numbers showing up will have sums between two and twelve, inclusive. So, in this case, n(S) = 36. Suppose you wish to determine the probability that you will roll a seven on a single throw of these dice. There are exactly six ways in which this can happen. If E = "the sum of the numbers is seven," then n(E) = 6. The probability of E, P(E), is simply determined by dividing n(E) by n(S), which in this example is P(E) = 6/36 = 1/6. So, aleatory probabilities are always determined by dividing n(E) by n(S), whatever E and S are, as long as E is a subset of S.

1.3.1.2 Relative Frequency and Statistics

Another way of assessing probabilities involves the many situations in which aleatory ground rules will not apply, but empirical methods are at hand to estimate probabilities. These situations arise when we have *replicable* or *repeatable* processes in which we can count the number of times events have occurred in the past. Suppose that, employing a defensible method for gathering information about the number of times event E has occurred, we determine the *relative frequency* of an occurrence of E by counting the number of times E has occurred, n(E), and then dividing this number by N, where N is the number of observations we have made, or the sample size we have taken. In this case, the relative frequency of E, f(E), equals n(E)/N. You recognize that this is a *statistical process* that can be performed in many situations, provided that we assume processes that are replicable or repeatable. It is true, of course, that a relative frequency f(E) is just an estimate of the true probability of E, P(E). The reason, of course, is that the number N of observations we have made is always less than the total number of observations that could be made. In some cases, there may be an infinite number of possible observations. If you have had a course in probability theory, you will remember that there are several formal statements, called the *laws of large numbers*, for showing how f(E) approaches P(E) when N is made larger and larger.

Probability theory presents an interesting paradox. It has a very long history but a very short past. There is abundant evidence that people as far back as Paleolithic times used objects resembling dice either for gambling or, more likely, to foretell the future (David, 1962). But attempts to calculate probabilities date back only to the 1600s, and the first attempt to develop a theory of mathematical probability dates back only to 1933 in the work of A. N. Kolmogorov (1933). Kolmogorov was the first to put probability on an axiomatic basis. The three basic axioms he proposed are the following ones:

Axiom 1: *For any event E, $P(E) \geq 0$.*
Axiom 2: *If an event is sure or certain to occur, which we label S, P(S) = 1.0.*
Axiom 3: *If two events, E and F, cannot occur together, or are mutually exclusive, the probability that one or the other of these events occurring is the sum of their separate probabilities. In symbols, P(E or F) = P(E) + P(F).*

All Axiom 1 says is that probabilities are never negative. Axioms 1 and 2, taken together, mean that probabilities are numbers between 0 and 1. An event having 0 probability is commonly called an “impossible event.” Axiom 3 is called the *additivity* axiom, and it holds for any number of mutually exclusive events.

Certain transformations of Kolmogorov’s probabilities are entirely permissible and are often used. One common form involves *odds*. The odds of event E occurring to its not occurring, which we label Odds(E, ¬E), is determined by Odds(E, ¬E) = P(E)/(1 – P(E)). For any two mutually exclusive events E and F, the odds of E to F, Odds(E, F), are given by Odds (E, F) = P(E)/P(F). Numerical odds scales range from zero to an unlimited upper value.

What is very interesting, but not always recognized, is that Kolmogorov had only enumerative probability in mind when he settled on the preceding three axioms. He makes this clear in his 1933 book and in his later writings (Kolmogorov, 1969). It is easily shown that both aleatory probabilities and relative frequencies obey these three axioms. But Kolmogorov went an important step further in defining conditional probabilities that are necessary to show how the probability of an event may change as we learn new information. He defined the probability of event E, given or conditional upon some other event F, as P(E given F) = P(E and F)/P(F), assuming that P(F) is not zero. P(E given F) is also written as P(E|F). He chose this particular definition since conditional probabilities, so defined, will also obey the three axioms just mentioned. In other words, we do not need any new axioms for conditional probabilities.

Now comes a very important concept you may have heard about. It is called *Bayes’ rule* and results directly from applying the definition of the conditional probability. From P(E* and H) = P(H and E*), you obtain P(E*|H) P(H) = P(H|E*)P(E*). This can then be written as shown in Figure 1.5.

This rule is named after the English clergyman, the Reverend Thomas Bayes (1702–1761), who first saw the essentials of a rule for revising probabilities of hypotheses, based on new evidence (Dale, 2003). He had written a paper describing his derivation and use of this rule but he never published it; this paper was found in his desk after he died in 1761 by Richard